

Making nuclear power usable again

The conditions that must be satisfied if nuclear power is to play a part in heading off the threat of global warming are easily stated, but require governments to swallow a little of their sovereign pride.

WHAT is the future for nuclear power, in Western Europe and elsewhere in the world? Several recent developments argue for a revival of the fortunes of a now neglected industry. The most pressing is the likelihood (and the hope) that in the decades ahead there will be increasingly tight restrictions on the quantities of fossil fuel that can be burned, on account of the carbon dioxide thereby dumped in the atmosphere. It is virtually impossible to see how major industrial countries will be able to meet the conditions likely to be required of them except by generating a substantial part of their electricity from nuclear fission. But there are many places, Japan for example, where there are persistent doubts about the security of supply of conventional fuels (whence the basis of Japan's considerable nuclear programme). Nuclear power has been in the doghouse since the Chernobyl accident in 1986. It is now a practical question to know how that state of affairs can be remedied.

First, suppress the temptation to declare that the Chernobyl accident was a consequence of such irresponsible behaviour by the operators of the reactor that caught fire that it could not occur elsewhere. The notion that people might casually carry out an unplanned experiment with an operating reactor, even one whose design does not ensure that its reactivity will decline as the temperature increases, cannot be ruled out for the rest of time. Even the well-drilled and well-supervised company of airline pilots is occasionally startled by gross departures from the rules: the case last year of a Russian pilot who put his young son disastrously in charge of an aircraft is a vivid illustration of that. Yet Chernobyl, whose consequences for the populations of the Ukraine, Belarus and Russia itself will become apparent only over the decades ahead, could easily have been even more damaging. In the wake of that disaster, civil nuclear powers signed agreements sponsored by the International Atomic Energy Agency (IAEA) acknowledging their cross-border obligations, but there is now a case for dusting these off. At the very least, reactor designs need to win the approval of people other than their national safety inspectors. Reactor operators should be licensed and supervised internationally. Erosions of national sovereignty on such a scale would be a price worth paying for security.

Highly active waste

Several other safety issues need attention, of which the most conspicuously contentious is waste disposal, which is a problem in two parts. The common belief that there is now no

means of storing highly active waste is mistaken. The process of embodying highly active waste in blocks of glass themselves contained in canisters of stainless steel is manageable and, if properly managed, should pose no threat to people or the environment. But what happens if, over centuries, stockpiles of vitrified waste are neglected? That cannot, of course, be allowed to happen. Accordingly, there is again a need for international supervision of national practices. That political price is likely to seem more serious to many governments than the economic cost of storage, but it is one that governments seeking to use nuclear power to meet the commitments on greenhouse emissions should be prepared to pay.

Other waste

Less than highly active waste is a different matter. Disposal of the waste materials arising in the normal operation of nuclear plants is impeded by the unwillingness of local and regional communities to see this material buried in their backyards. The arguments put up do little credit to those who make them, but they are politically compelling in democracies (and other kinds of states appear to have been dangerously casual in their management of waste disposal). It is difficult to see these problems being solved unless governments individually are prepared to invest some of their political capital in winning consent to sensible proposals. The way in which the British government has weakly let the issue of waste disposal hang fire over the past several years is an illustration of how not to conduct these affairs — and a measure of its own poor stock of political capital.

There remains the issue of reprocessing. Should the plutonium inevitably generated in reactors operating with uranium be extracted, or left with the spent uranium from which it springs? The case for reprocessing is that plutonium is itself a fuel (and can already be used in reactors using 'mixed-oxide' fuel, without waiting for fast reactors to be proved safe) and that plutonium can be stored more safely if unencumbered by the much more radioactive fission products generated in reactors. The case against is that separation is costly, and that plutonium can be used to make nuclear bombs, at least if the fissile isotope (^{239}Pu) is not too heavily contaminated by isotopes fissile only with fast neutrons. The arguments on either side of this question, almost always coloured by the mythical status acquired by plutonium, have been pre-empted by the distaste of the US government (since President Jimmy Carter's time) for



reprocessing not only domestically, but by countries to which the United States ships uranium.

It is high time that this question was reopened. Safeguarding plutonium is important. Given that the element will not now disappear from the surface of the Earth, it is also a perpetual obligation. There seems no sensible alternative to a scheme whereby existing stocks (in weapons as well as non-weapons states) are physically transferred to some international stockpile where their integrity can be ensured. That notion, widely canvassed in the 1970s, should be revived. At the same time, governments (and the US government in particular) could help to take the edge off worries about bombs made with illicit plutonium by disclosing what they know about the purity of the plutonium required to make a bomb. The US nuclear explosion in 1961 of a device fashioned from 'reactor plutonium' was, in the event, carried out with material supplied by the British government from British reactors in exchange for enriched uranium. Given the design of the British reactors then operating, this material is likely to have been more concentrated in ^{239}Pu than that from oxide fuels. Saying so would be in everybody's interests.

What all this implies is that there are several steps that must be taken, by governments alone and in concert, if nuclear power is once more to be an accessible source of energy. Unsurprisingly, many of them require international arrangements that blunt the edge of national sovereignty. If nuclear power is potentially part of the means by which the threat of global warming will be avoided, who can protest that global conditions must be satisfied if the expectation is to become reality? Many, of course, will wring their hands at the prospect that meeting these conditions will entail an abridgement of national sovereignty. But they have no choice. The unique feature of the global warming business is that its global character enforces mutual interdependence, not only in the arrangements for making nuclear power usable, but also in making the emission of greenhouse gases in one state a proper concern of those who live elsewhere. But if the future requires that governments should put up with external scrutiny of their carbon-dioxide emissions, is it not even more important that they should allow that their nuclear plants should be supervised internationally, and that their plutonium should be stored in common repositories? The sooner that is recognized, the better for us all. □

History and the lotteries

Britain should not have bought the Churchill archives without ensuring that full use can be made of them.

The British National Lottery seems destined for a bumpy ride. After less than half a year, it has discovered that winners are ambivalent, to say the best of it, about the publicity attending their sudden conversion into millionaires, that the first sob stories have come to light about people who are so addicted to the new game that they spend all they have on

it, neglecting the necessities of life, and a scheme for providing gamblers with instant satisfaction (or despair) has been found easily corruptible. But is it not all in a good cause? For are not roughly half the proceeds of the lottery to be devoted to good causes — marking the impending millennium, preserving national monuments, assisting charitable causes (whose normal income has fallen drastically because of competition from the gambling machine)? Sadly, even that aspect of this plan is turning sour.

Last week, the body known as the Heritage Commission decided to spend some of its first lottery receipts on buying from the trustees of the family of Sir Winston Churchill, Britain's wartime prime minister, the paper archive left by the grand old man. By all accounts, the sum involved is £12.5 million. Hitherto, the archive has been lodged at Churchill College, Cambridge, where it has been carefully preserved and partially catalogued, mostly with the use of public money. By all accounts, a substantial part of the collection includes state documents, generated by Churchill during his long spell in public office. There has been a great fuss. Expectations that the purchase would seem apt just ahead of the fiftieth anniversary of the ending of the Second World War in Europe have been disappointed. Winston Churchill's grandson, a Member of Parliament who carries the same names, has been accused of greed.

But this complaint, and its accuracy, is a sideshow except in that it illustrates that the financial imprudence engendered by gambling infects not only the gamblers but the beneficiaries. A much more serious issue stems from a little-noticed condition of the sale of the Churchill archive — the understanding that while the physical archive has been sold, copyright in the material it contains will continue to be vested in the family trustees. That is an iniquitous condition, if a fairly common one. It is standard practice, when people sell their papers to, say, a university library that they or their heirs retain the copyright. Up to a point the condition is justifiable. It would, for example, be unjust that a published work by an author could be republished by any publisher with the inclination simply because the manuscript is placed in a public archive. The iniquity in these arrangements is that unpublished writings enjoy a much stricter form of copyright protection: they cannot be published without the consent of the copyright owners, who may levy a charge.

That practice largely defeats the purpose of making archives public property. Scholars can spend days reading through old papers to form a view about some sequence of events, and then discover that the copyright owner will not let the evidence be published. Often, consent is withheld because the owners fear that an ancestor or other relative will be shown in a bad light. The law (everywhere applicable) does not require that copyright owners have a duty to history, which is proper when an archive is privately held. But that should not be the case when, as in the case of the Churchill papers, money has changed hands and the cost of conservation and access put on the shoulders of taxpayers. Those now complaining about the younger Churchill's gambling windfall would better direct their attention to this issue of principle. □

Project to sequence human genome moves on to the starting blocks

London. Almost exactly ten years after the idea was first suggested, funding agencies in the United States and Britain have given the green light to pilot projects designed — if successful — to lead to the first full-length sequence of the human genome.

Next Saturday (13 May), John Sulston, director of the Sanger Centre in Britain, and Bob Waterston, director of the Genome Sequencing Center at Washington University, St Louis, plan to outline a joint proposal for an international effort to achieve this objective to a meeting at Cold Spring Harbor, New York.

Research teams at various centres would be responsible for helping to sequence a set of cosmid covering each chromosome in turn. These sequences would then be combined, and the complete sequence data would eventually be placed in the public domain.

Although the specific strategy being suggested by Sulston and Waterston would, for technical reasons, result in a sequence that still contained some gaps, many feel that its cost — estimated at between \$300 and \$500 million — is sufficiently lower than earlier estimates to compensate for any omissions.

Sulston is also confident that a fully mapped sequence, even if only 99.9 per cent accurate, would still be invaluable to the whole research community. He says that the sequence could be established in as little as five years, and that the task is now timely and urgent. "Every year that goes by without embarking on this, we are wasting both money and time."

Agencies such as the National Institutes of Health (NIH) and the Department of

Energy in the United States, and the Medical Research Centre (MRC) and the Wellcome Trust — the main supporter of the Sanger Centre — in Britain, appear to be in agreement.

The NIH, for example, has issued a request for applications for pilot large-scale sequencing projects, and is prepared to give applications for any such project "expedited treatment" in order for it to start by the end of the year. It has earmarked \$20 million to be distributed between such projects and work on advanced sequencing techniques.

Similarly the MRC has decided to make available some of the extra money it won from the government in this year's science budget for 'strategic' objectives.

Both Sulston and Waterston say they intend to apply for the funds being made available. Whether their specific proposals receive approval will depend on how they fare during the peer review process.

But, given the strong track record of both groups in sequencing the genome of the nematode worm *Caenorhabditis elegans*, there will be little surprise if their applications are successful.

Whatever the outcome, there is little doubt that the joint proposal has helped to crystallize a growing consensus that, after ten years of preliminary work, the time has arrived for a preliminary assault on the 3-billion bases of the complete human genome.

"There is a lot of interest in moving on to sequencing more quickly," says Elke Jordan of the NIH's National Center for Human Genome Research (NCHGR). "In the past few months, we have gone from a feeling that we were not quite ready to a feeling that 'yes, we should go'."

There is similar enthusiasm at the Wellcome Trust. "We have not yet agreed to fund the project; that will depend on the outcome of the normal peer review processes," says Michael Morgan, programme director at Wellcome. "But we would not be in this position [of accepting proposals] if we did not accept that it is a sensible thing to do."

Up to now, support for genome research by the NCHGR and other bodies has been distributed relatively widely. One priority, for example, has been novel sequencing technologies, based on the argument that large-scale sequencing should wait until the costs become significantly lower.

Some, indeed, continue to support this point of view, and advise caution before rushing ahead with the strategy that Sulston and Waterston are suggesting. Others, however, point out that there have been few

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Worm's eye view: sequencing techniques for the nematode genome may now be applied to humans

major technical advances in sequencing in the past couple of years — and that waiting longer could delay the project unnecessarily.

In contrast, the proposal being put forward by Sulston and Waterston is built largely on techniques already in use in both laboratories to sequence the nematode as well as parts of the human genome.

The approach they suggest would be based on the so-called 'shotgun' sequencing of a minimal set of cosmid covering each human chromosome in turn. The resulting sequences will then be put together, with additional effort applied to 'finishing' the data where necessary.

Sulston says that, on the basis of their previous experience and simulations with available data, he and Waterston believe that the resulting sequences will contain no more than 3 gaps per cosmid — caused chiefly by long tandem repeats, inverted repeats and homopolymeric stretches. Nevertheless the sequence would still contain sufficient information to act as a fine-grain map containing, for example, the accurate location of the entire set of human genes.

"The general idea is to make a thorough attempt at maximal representation of the genome both at the clone level and at the sequencing level, and to tie the whole thing together by identifying the landmarks," says Sulston. "This will provide a means of organizing as much as possible of the information that is already being developed."

Sulston and Waterston claim that, if carried out on a large enough scale, the cost of sequencing could be reduced to about 10 cents a base (for comparison, the nematode ▶

Technicians die in Ariane V accident

Paris. The European Space Agency (ESA) has set up an inquiry into the deaths last week of two technicians from France's national space agency (CNES) during pre-test verification checks on the cryogenic engines for the launcher Ariane V.

The technicians died by asphyxiation following the leak of liquid nitrogen. The inquiry is expected to provide details on the origins of the leak, which took place during inspection of the umbilical mast used to fuel the launcher on the launch-pad.

A spokesman for ESA said that the accident is unlikely to delay the first launch of Ariane V, still scheduled for the end of November, with the scientific mission Cluster as its pay-load. □

sequence, which is now 15 per cent complete, and is expected to be finished by the end of 1998, costs about 50 cents a base.)

Some remain sceptical about whether the figure can indeed be reduced this low. "We have yet to see anything on paper to allow us to judge whether it is feasible," says one sequencer while admitting that, in principle, he is strongly supportive of the approach that Sulston and Waterston are taking.

Even at this low price, however, Sulston is keen to emphasize that, for scientific and other reasons, he would like to see the project carried out collaboratively between a number of centres, with different research groups agreeing to concentrate on different chromosomes.

Such an approach is likely to go some way to meeting fears that arose earlier this year, when the proposals were first being discussed by genome research centres, that the whole project would be dominated by the Sanger Centre and the St Louis group. "We have each said that we would be willing to do one third of the operation, but we do not insist on doing it all; the intention is to have a fully cooperating network."

Cooperation — not competition (apart from that for funding, with peer review ensuring the quality of the data produced) — will be the key. "This is a place where the market does not work well," says Sulston. "It is inappropriate when you are trying to establish a framework where there is only one solution; you need a strategic plan."

David Bentley, head of human genetics at the Sanger Centre, suggests as a possible model the consortium of laboratories set up in Europe, with the backing of the European Commission in Brussels, to sequence the yeast genome. "It has been put together in a very informal and flexible way, and people have shifted their aspirations according to what they have been able to achieve," says Bentley. "We will try to work in the same way with the human genome."

If the consensus behind the overall strategy is still growing, it is not yet complete. Some still argue, for example, that a more effective approach would be to build on the knowledge of detailed sequences being established through cDNA libraries, in other words starting with a focus on functioning genes. Others continue to prefer a strategy known as 'skimming' the genome, carrying out the complete sequencing of randomly selected sections.

The Cold Spring Harbor meeting is expected to provide a forum for vigorous debate between those supporting rival strategies. But for Sulston, Waterston and their colleagues, the funding agencies prepared to back their approach, and industrial researchers keen to exploit their results, the main question now is not whether, but when, to start work. If the proposed pilot projects succeed, the main sequencing would be ready to start by the end of the next year. And the whole project could be finished by the year 2001 — five years earlier than many had previously estimated.

US court rules discovery of gene sequence 'not obvious'

London. In a major boost to the biotechnology industry, a US appeals court has ruled that the discovery of a novel gene sequence cannot be described as 'obvious' — and that the sequence in question can therefore be legitimately included in a patent — even if the existence of the gene was previously known, as were techniques for obtaining its complementary DNA (cDNA).

The ruling is based on the fact that, even though the gene can be deduced from the protein that it expresses, so-called 'redundancy' in the genetic code means that its precise nucleotide sequence can be considered unknown until it has been 'discovered'. (The redundancy results from the fact that more than one nucleotide base triplet in a DNA sequence can code for the same amino acid in the protein, and that a wide range of sequences can therefore theoretically code for the same protein.)

The court's decision is a direct challenge to those who have been arguing that the controversy over patents on human and animal genes is likely to blow over relatively quickly as sequencing and cloning techniques become routine laboratory practice, and the discovery of new gene sequences, therefore, a relatively unremarkable affair.

Indeed, one of the main implications of the ruling is to limit the range of techniques whose existence can be invoked to dismiss a discovery as 'obvious', and therefore to increase the degree of predictability in the US patents system.

At the same time, the ruling is being seen in the biotechnology industry as putting gene sequences on the same level — at least as far as patent protection is concerned — as chemical molecules, whose eligibility for patent protection (provided they fulfil criteria of novelty and usefulness) is clear.

"The implication is that new technologies in the biotechnology and genomics area can be covered by existing patent laws in the same way as other areas such as chemistry and physics," says Robert Benson, attorney for Human Genome Sciences (HGS) Inc. in Rockville, Maryland. HGS has over 70 patent applications on partial and full gene sequences awaiting a ruling from the US Patents and Trademark Office (PTO), and hopes the court ruling will increase their chances of being approved.

The decisions concern the application by Thomas F. Deuel and three fellow scientists at Washington University in St Louis, Missouri, for a patent on isolated and purified DNA and cDNA molecules encoding heparin-binding growth factors (HBGFs), proteins that stimulate cell division and therefore facilitate the repair or replacement of damaged or diseased tissue.

A PTO examiner had initially rejected the patent application — which includes coverage on all possible DNA molecules coding for the disclosed proteins — on the grounds that cloning the gene would have been "obvious" to anyone "of ordinary skill in the art at the time of the invention".

The examiner quoted two instances of 'prior art' to support this view. One was that a heparin-binding protein had already been identified and isolated from human and bovine brain tissue, and both proteins had been found to share the first 19 amino acids in their N-terminal sequences. The second was that methods for isolating either DNAs or cDNAs by screening sequence libraries had been initially described by Tom Maniatis of Harvard University in 1982.

The examiner argued that it would have been relatively straightforward to use the known N-terminal sequence, and the techniques described by Maniatis, to design a probe to isolate a gene encoding the protein from a cDNA library. Although challenged by the applicants, the ruling was subsequently upheld in November 1993 by the PTO's own appeals board.

Last month, however, in response to a challenge mounted with the backing of both the Biotechnology Industry Association and the Bay Area Science Centre in San Francisco, as well as biotechnology companies such as Chiron Inc., the US Federal Circuit court overturned the PTO's objection.

The court ruled that, even though knowledge about the proteins identified earlier meant the general chemical nature of the cDNA molecules in the patent claims "may have been obvious" — and the knowledge that "some gene existed" may have been clear — redundancy in the genetic code meant that the precise cDNA molecules isolated from the library "would not have been obvious".

"Until the claimed molecules were actually isolated and purified, it would have been highly unlikely to one of ordinary skill in the art to contemplate what was ultimately obtained," said the court. "What cannot be contemplated or conceived cannot be obvious."

More broadly, the court stated that a general motivation to search for some gene that exists "does not necessarily make obvious a specifically-defined gene that is subsequently obtained as a result of that search".

It added: "Even if, [. . .] the existence of general cloning techniques, coupled with knowledge of a protein's structure, might have provided motivation to prepare a cDNA or made it obvious to prepare a cDNA, that does not necessarily make obvious a particular claimed cDNA". **D. D.**

Split threatens young scientists' group

Washington. A highly public feud between founder members of the Young Scientists' Network (YSN) and some of its more radical participants is threatening to split the body responsible for a pioneering computer bulletin board which has served as an important focal point for young researchers facing cut-throat competition for research posts.

Relations between the two factions have worsened since the first group inadvertently chose 19 April — the day of the Oklahoma bombing — to accuse the second of 'infoterrorism' in distributing leaflets to high school students attending a meeting on 13 April at the University of Texas at Dallas carrying information about the difficulty researchers face finding jobs.

Some prominent members of the network, including Kevin Aylesworth, who founded YSN in 1990 and has since become active in science policy circles in Washing-

ton, are now considering setting up a formal organization to represent young scientists on a far broader front.

Aylesworth and his allies believe that the issue which the network once addressed — that young people were being misled into thinking there was a shortage of scientists, while in fact the United States was producing a glut of PhDs — has now been resolved. They now favour a shift of emphasis towards defending science from impending Congressional cuts.

On 28 April, a message was posted on the bulletin board signed by Aylesworth, John Sahr and John Quakenbush — both administrators of YSN — and five others, saying that "the war of 'the Myth' [of a scientist shortage] has, to a large extent, been won." It called on YSN members to mobilize instead against Congressional cuts.

But the proposal to redirect the activities

of the network drew scorn from some of the 3,000 young scientists who subscribe to it. Early response to the posting was largely negative. "A few national magazine and newspaper articles in a four year period do not a war win," wrote Ray Greene from Cornell University, New York.

The main point at issue between the two groups appears to be whether the YSN has, in fact, met its original objective by effectively dispelling the so-called myth of a scientist-shortage. Aylesworth says that it has. Today, he says, "no-one on Capitol Hill believes that there is a scientist shortage".

Others on YSN — such as Gene Nelson, an unemployed biophysics PhD living in Dallas who did the leafleting at the Texas meeting, which was addressed by Nobel Laureate Francis Crick — think the myth is alive and kicking among the general public. In a recent posting, Nelson drew a parallel between young people going into the performing arts and science. "My strong advice now is, don't give up the day job," he wrote.

Supporters of Nelson point out that he was merely distributing publicly-available information to high-school students, who should be old enough to assess it before making career decisions. But his opponents think that advising school children to avoid science is irresponsible, and that support for such an approach will alienate official science lobbying groups such as the American Physical Society (APS) with which young scientists need to work.

The core of the dispute appears to be whether the young scientists ultimately share the same interests as their older colleagues. Aylesworth thinks they do. There is now "a pretty big gulf" inside YSN, Aylesworth says. "Eventually, some new organization may emerge."

For Nelson, the network exists to snap at the heels of APS and other bodies — not join them at the tables of power. He says there is no need to split. "The last thing the YSN needs in these perilous times is to splinter apart," he says. **Colin Macilwain**

Clean-up funds cut

Washington. The US Department of Energy, Washington State and the Environmental Protection Agency have agreed \$1 billion of cuts in the clean-up programme at Hanford, the United States' most problematic nuclear site, over the next two years.

Tom Grumbly, assistant energy secretary, said last week that the cuts, which represent one-quarter of the Hanford clean-up budget, would come from competitive tendering and reduced overheads and paperwork. But with the threat of even deeper cuts from Congress, environmentalists are worried that the already glacial progress on clean-up at Hanford may come to a halt. **C.M.**

Director quits over A-bomb exhibit

Washington. Martin Harwit, the director of the National Air and Space Museum (NASM) in Washington DC has resigned, three months after the Smithsonian Institution abruptly cancelled a controversial exhibition there intended to commemorate the fiftieth anniversary of the dropping of the atomic bomb and the end of the Second World War.

The NASM is one of several museums and galleries in Washington that come under the aegis of the Smithsonian. Harwit's departure had long been sought by the Air Force Association and other critics



Under siege: National Air and Space Museum.

of the cancelled exhibition, which was to have used as a focus the aircraft that dropped the bomb, Enola Gay.

But according to both staff and outside observers, it leaves curators at the embattled institution more anxious than ever about its ability to maintain its integrity in the face of external political pressure.

"The curators are terribly discouraged and demoralized," says Stanley Goldberg, a consultant to the Smithsonian and a leading historian of science, who served on the advisory board of the exhibition. "The question is whether scholars can do their work without having to make sure that it agrees with the prejudices of outsiders looking over their shoulders".

Harwit, an astrophysicist, claims that his departure is in the best interests of the museum which, with more than 8 mil-

lion visitors last year, is the most popular in the United States. "I had to decide whether staying on was helpful or detrimental to the museum," he says. "I am happy that the controversy will die now. People will be satisfied, and the museum will be able to take a new direction."

But some of Harwit's former colleagues still at the museum are

somewhat less sanguine. "We're in limbo," says one of Harwit's former colleagues. "We don't know what our new direction will be, or what our mandate will be." A colleague at the

museum adds: "The situation is clearly very bad, and the resignation is just the latest chapter. It is a victory for our most unreasonable critics."

These critics can expect a field day today — 11 May — when they testify before the powerful Senate Rules committee, chaired by Ted Stevens (Republican, Alaska), at a hearing on the performance of the Smithsonian.

In a statement issued after Harwit's resignation, Michael Heyman, the institution's secretary, praised his director's eight-year record. But the statement said little about the Enola Gay fracas, and Heyman declined to comment further. Asked if Heyman was the right man to lead the Smithsonian at this time, Harwit said that after six months, it is "much too early to judge". **C.M.**

UK academics see research income rise

London. Research in British universities appears to be booming, according to the publication last week of statistics showing that the research income per full-time academic scientist has trebled in cash terms over the past ten years.

But in a statement accompanying the announcement, the Committee of Vice Chancellors and Principals (CVCP), said the rise in research income from £12,000 (US\$19,200) per academic to £36,000 coincided with a fall in public investment.

Furthermore, according to the 1995 edition of the CVCP's *University Management Statistics and Performance Indicators in the UK*, expenditure on university libraries, if taken as an indicator of infrastructure investment, fell by 22 per cent in real terms in the decade between 1984 and 1994.

The CVCP survey does not indicate which sources provided most research income. But a separate report published last month by the Universities Statistical Record shows that funding from the European Union and from UK charities had risen the most, by 27 and 18 per cent respectively, over the past year; in contrast, spending by UK industry rose by only 7 per cent, and by the government by 5 per cent.

The Association of University Teachers,

the academics' labour union, has given a mixed reaction to the statistics. Monica Hicks, a spokeswoman for the union, acknowledges that diverse sources of income free universities from the "whims of government". But she describes the fall in public investment as "a national disgrace". Investment in universities, says Hicks, "is the responsibility of government", not least to protect 'blue skies' research.

Out of the 37 broad subject bands covered by this year's CVCP survey, biochemistry has replaced clinical medicine as the field attracting the most money per head. Research income per full-time biochemistry academic was found to be just under £96,000, compared to £74,000 in 1991-92.

The University of Leicester has overtaken Dundee as the most productive venue for biochemistry research. Leicester's income per full-time academic jumped from £123,000 in 1991-92 to £237,000 in 1993-94. The University of Oxford came second with £214,000, and Imperial College, London, third at £162,000.

Materials science, mineral engineering and physics were the other strong performers, attracting £94,000, £82,000 and £76,000 respectively. In the first category, Queen Mary and Westfield College, University of

London, and the University of Birmingham were neck-and-neck, each attracting over £191,000 per academic.

Heriot-Watt University (£280,000) continues to dominate mineral engineering, while University College London (£151,000) overtook the University of Cambridge (£149,000) to head the physics table.

The section on graduate unemployment shows that while there are fewer long and short-term jobless science graduates, unemployment continues to rise among zoology, botany, chemical engineering, mineral technology and genetics graduates.

Unemployment of recent genetics graduates rose from 16.8 per cent in 1992-93 to 17.9 per cent in 1993-94; while the number of out-of-work biotechnologists, having fallen from 21.6 per cent to 10.3 per cent in 1992-93, rose to 25.8 per cent.

The CVCP statistics were compiled from information received from government research councils, government ministries, industry, overseas sources and charities. The figures apply only to the United Kingdom's 56 'old' universities. Next year's statistics will include information from all of the former polytechnics that were accorded university status under the 1992 Further and Higher Education Act.

Ehsan Masood

Europe approves energy funding — but blocks projects

Paris. For many scientists, getting a research programme approved can be easy compared with obtaining the requested funds. The European Union (EU) has now created the opposite problem. It has approved the spending by the European Commission of ECU30 million (US\$39 million) on a programme on renewable energy, while at the same time refusing to approve the programme itself.

The dispute has arisen over the future of part of THERMIE, a research programme that ran from 1990 to 1994. One part is assured annual funding of ECU100 million, as it is part of the ECU967-million 'non-nuclear energy' section of the fourth Framework programme, which lasts from 1994 to 1998. It was considered eligible because it funds 'demonstration projects' of technologies considered too young and risky for industry to finance.

But the commission has ruled that the remaining part of THERMIE, which deals with testing demonstration projects under different economic and geographical conditions, cannot legally be included in Framework because it is concerned with economic rather than technological risks.

To meet this problem, the commission has proposed a separate ECU30-million-a-year programme, confusingly called THERMIE II. Both Germany and the United

Kingdom, however, oppose this proposal, arguing that they agreed to the overall envelope for renewable energy research when they signed up for the Framework budget — and are not prepared to pay a penny more.

Officials in the UK Department of Trade and Industry add that both countries also

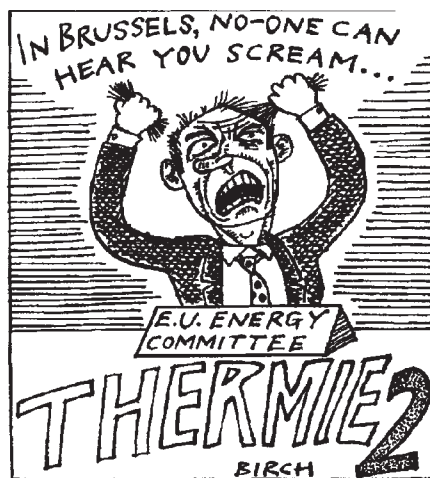
ministers — the body through which member states are represented directly in Brussels — earlier this year. The 13 member states in favour of the proposal were able to push through its budget on a majority vote; but Germany and the United Kingdom were able to block adoption of the programme, as this requires unanimity.

Last month, the energy committee of the European Parliament condemned the council for "endangering" research centres across Europe. As the parliament has no powers in this area to contest the council decision formally, it issued a resolution calling on council to "recognize the need to adopt the programme".

The result of the stalemate is that ECU30 million is now sitting in a bank, earmarked for renewable energy research; but not a penny can be spent. In an attempt to resolve the issue, Christos Papoutsis, the energy commissioner, has suggested an informal tripartite meeting between the council, the commission and parliament.

The issue will probably be discussed at a meeting of the council of energy ministers on 1 June. By then, a solution may be even more elusive; France, which initially opposed the proposal but had a change of heart during its current six-month term as president of the EU, may have reverted to its initial position.

Declan Butler



oppose the proposal as being too near-market, and claim it could set a precedent of using EU research funds as industrial subsidies, thus distorting the market.

The issue took on an air of absurdity when the dispute reached the council of

Germany refuses to prolong university support scheme

Munich. Germany's federal government last week announced that it will not extend a programme set up six years ago to help universities through an overcrowding crisis generated by the 'baby boom' of the 1960s.

But the announcement has prompted an angry reaction from the University Rectors Conference (HRK). By abandoning the programme, says the HRK, the government could be responsible for the loss of thousands of university jobs at a time when student numbers are rising at a record rate.

Germany's post-war constitution gives full responsibility for education to the *Länder*, to avoid an excessive concentration of power over education. Nevertheless, the federal government contributes half of universities' building costs, and provides much of the support for university research.

The system worked well until massive rises in student numbers, combined with the more recent recession, put an increasing strain on resources. Overcrowded classrooms have resulted primarily from a constitutional right guaranteeing a university education to all students with the school leaving certificate, the *Abitur*.

The federal government believes that universities should help themselves by compromising on this right and restricting student numbers in over-subscribed subjects. It has, however, been willing to help them meet the impact of the baby boom.

In the late 1980s, the government agreed to provide financing towards a DM2.1-billion programme to ease the pressure of student numbers in heavily subscribed subjects such as informatics. Known as the first *Hochschulsonderprogramm* (HSP1), this created 3,200 positions nationwide, but runs out at the end of this year.

The HSP1 was supplemented in 1991 by an additional 10-year DM4-billion programme, HSP2, which had four broad aims: to create new positions for women, for young scientists, and within the technically orientated institutes of higher education (*Fachhochschulen*), and to support European collaboration in the universities. The 60 per cent federal share of HSP2 funding is

higher than in HSP1, where financing was balanced equally with the *Länder*.

Both programmes support only west German universities, and a third, DM2.4-billion programme (HEP) was therefore set up for eastern German universities. Because of their special problems, the federal government paid three-quarters of this bill.

The federal and *Länder* governments have now agreed that the HSP2 and HEP programmes should be extended beyond 1996, so that general trends — such as the continuing low number of women in academic life — can be influenced centrally.

They also agree that the two programmes should be merged. How costs will be shared is not yet known, but the *Länder* will probably challenge the federal government's suggestion of a 50:50 split.

More controversially, the federal government is to abandon the HSP1. "The federal government will not allow any claims in areas in which the *Länder* carry the responsibility, be it political or financial", says Berndt Neumann, parliamentary secretary of state for education and research. Curricula come into this category, he argues.

In a statement two weeks ago, Neumann pointed out that the *Länder* had not made any efforts to reduce student numbers by introducing limits, despite agreeing two years ago to consider such a move.

The universities see the federal government's attitude as potentially damaging. Gerhard Neuweiler, professor of zoology at the Technical University in Munich and a former head of the German science council, the Wissenschaftsrat, says that as the government sets national priorities in science, such as informatics, it must help foot the bill because most *Länder* cannot afford to.

Disillusion is widespread in universities. Last month, Adolf Theis resigned as rector of the University of Tübingen in protest at the lack of support by the federal government for universities. If this does not change, he warns, politicians will face serious conflict. "Tensions among students is rising in a manner reminiscent of 1968."

Alison Abbott

CERN makes good on technician's sabotage

London. The Large Electron Positron Collider (LEP) at the European Laboratory for Particle Physics (CERN) was switched on again last week after a four-month maintenance shut-down — despite the sabotage in February of two of its particle accelerators.

Fears were raised over the loss of laboratory time following the discovery on 13 February that one of CERN's long-serving technicians had cut about 200 cables and

hidden 1,300 electronic components needed to control the proton-synchrotron accelerator and booster (see *Nature* 373, 650; 1995).

But CERN's antiproton programme, which started on 21 April, one week later than planned, appears to have been the only casualty. The collider itself restarted on time, after the laboratory's technicians and engineers had worked "flat out" to put right the damage caused. □

UNCTAD to help former Soviet scientists prosper

London. The United Nations agency responsible for trade issues is to set up a special programme to help scientists in the former Soviet republics to negotiate with foreign companies interested in licensing their research results.

The programme, to be launched by the United Nations Conference on Trade and Development (UNCTAD), will help research and development managers in the former republics to master international business skills and intellectual property protection.

The UNCTAD programme will be paid for from funds allocated to the agency for assistance for the states of the former Soviet Union by the United Nations Development Programme (UNDP) in New York. It was approved last week by the Commission on International Investment and Transnational Corporations.

The programme will offer technical assistance to governments to establish policies and laws supporting scientific development. It will also provide international expertise on issues such as trade legislation, how to identify potential foreign partners and negotiation skills.

UNCTAD officials claim the large amounts spent by the Soviet regime on science have produced valuable resources which, if tapped properly, could generate a significant income for the republics.

But at present, they say, many risk losing their scientific and technological resources as their nascent market economies do not have the marketing skills to make the best use of this investment. The UNCTAD programme aims to help halt this decline by offering a structure for marketing such scientific resources internationally.

Maurice Odile, a spokesman for UNCTAD, says that one of the difficulties is that many scientists are unaware of the potential value of their intellectual property, or how either to ensure its protection or to market it. Yet without such knowledge, UNCTAD fears that such scientists risk being exploited by foreign corporations.

Scientists previously operated in the security of the Soviet system which provided intellectual property protection by issuing so-called 'authors' certificates' of relatively low value, with the results becoming the property of the state.

With the collapse of communism, the flow of funding for science has slowed to a trickle in most of the former republics. UNCTAD started assisting countries such as Belarus, Estonia, Latvia, Lithuania, Ukraine and Uzbekistan in 1992, helping them to make strategic alliances and business deals with corporations.

Peter O'Neill

Airport plans threaten dolphin sanctuary

Hong Kong. The Hong Kong government has bowed to pressure from local green groups and announced it is to set up a 1,000-hectare marine sanctuary for the rare Chinese white dolphin (*Sousa chinensis*).

But campaigners are opposed to the government-funded HK\$18-million (US\$2.34 million) project, to be managed by the Agriculture and Fisheries Department (AFD). They argue that the future of the dolphins may be compromised by the siting within the sanctuary site of a temporary aviation fuel storage depot.

The Provisional Airports Authority (PAA), the government body set up in 1990 to plan, design and construct Hong Kong's new airport at Chek Lap Kok, was given approval in early April to begin construction of the Aviation Fuel Receiving Facility (AFRF) at Sha Chau, one of two islands to be included in the sanctuary. Fuel from the depot will be transferred to the airport platform by an under-sea pipeline.

Michael Lee, deputy director of the AFD, says that demarcation of the area will begin next year. But protesters say it should have been sited further away from an area used by the dolphins.

Joanna Ruxton, senior conservation officer for World Wide Fund for Nature Hong Kong, is convinced that a combination of human activity, the blasting during the construction of the fuel-receiving facility and the volume of shipping traffic to and from the site will be detrimental to the dolphins in an area that is their last stronghold. Their favoured habitat of shallow coastal waters is gradually disappearing as a result of development and dredging.

Other critics of the PAA include Lindsay Porter and Chris Parsons, two PhD researchers who are part of a HK\$2-million government-funded dolphin research project at the Hong Kong University Swire Institute of Marine Science.

But Richard Morse, environmental manager at PAA, argues that the authority has



No hiding place: shipping already endangers habitat.

made a "good faith effort" to choose the best site for the fuel depot out of 10 possibilities. Under the "exhaustive process" for selection, undertaken in consultation with 16 other government departments, Sha Chau emerged as the "only viable option" after all environmental considerations were taken into account.

The Chinese white dolphin or Indopacific humpback dolphin, is found in coastal and estuarine waters from the southern tip of Africa, around India, Malaysia, down to northern Australia and north into the South China Sea. The taxonomic status of the different populations is still being debated — mainly because of a lack of research data — but the population in Chinese waters is remarkable in that grey calves fade to a

striking pinky white as adults.

PAA says that independent experts have confirmed that a fuel facility based in Sha Chau is not likely to harm the dolphins. Bern Wursig, director of the marine mammal research programme at the Institute of Marine Life Sciences at Texas A & M University, was brought in by the authority to examine the potential effects of the fuel facility on the dolphin population.

While Wursig concluded in his report that "the AFRF at Sha Chau, by itself, is not likely to have negative effects on *Sousa* as a population", he did recommend several mitigating measures, including the setting up of "the largest possible" sanctuary for the dolphins and an overall Hong Kong dolphin management programme.

Observers are unhappy that Wursig's conclusions were based on just four days spent with the dolphins. The Hong Kong University researchers, on the other hand, have spent 18 months in the field. But Wursig argues that these mammals "habituate remarkably well" to noise and interference caused by humans.

As evidence of this, Wursig cites the reappearance of the pink dolphins around the Brothers, a group of islands that were virtually removed for the airport's construction. It is, however, difficult to establish whether the same numbers of dolphins returned as there are no original figures for comparison.

Despite the differences of opinion on whether a dolphin sanctuary can be sited near a fuel facility, Ruxton, Parsons and Wursig share common ground in considering pollution a threat to the dolphins. "Things going on noise-wise [for the dolphins] are not so horrible. But the things going on pollution-wise and habitat-degradation-wise are horrible," says Wursig.

All three researchers believe that pollution from untreated sewage, heavy metal toxins, toxic mud pits resulting from dredging operations, over-fishing, increased volume of shipping traffic and the rapid removal of the shallow coastal habitat preferred by the dolphins, all combine to constitute a major threat to the dolphins.

But the dolphin/fuel receiving facility issue may have become a victim of the island's delicate but often murky politics — the dolphin sanctuary may have been agreed in exchange for the fuel facility.

Nevertheless, Wursig wants the dolphins to benefit as much as possible. "The Provisional Airports Authority gave the sanctuary so that they could have their interim reserve facility," he says. "But it's now time for the green groups to say, okay, let's work together on this to put together the best programme for the dolphins." **Maggie Verrall**

Plans unveiled for giant ray detector

Washington. Construction work on what a group of physicists hope will eventually be the world's largest detector of cosmic rays could begin as early as 1997 — providing funding can be raised from US and foreign sources — according to one of the leaders of the design project.

James Cronin of the University of Chicago was speaking during a meeting last week at the Fermi National Accelerator Laboratory near Chicago (Fermilab). The meeting took place halfway through a six-month design study which members of the Pierre Auger Cosmic Ray Project hope will lead to the construction of two large detector arrays, one each in the northern and southern hemispheres.

The arrays would be used for studying the most energetic particles in the universe. The project, which originated in 1992, would involve two arrays, each using 3,000 garage-sized detectors spaced over an area of 5,000 square kilometres and located in a

remote region. The largest current cosmic ray detector, in Japan, is only 100 square kilometres in area.

But it will first be necessary to sign up international partners and commit funding. Cronin, a leader of the study project along with Alan Watson of Leeds University in the United Kingdom, estimates the costs at \$100 million to \$200 million for a full detector array in both hemispheres.

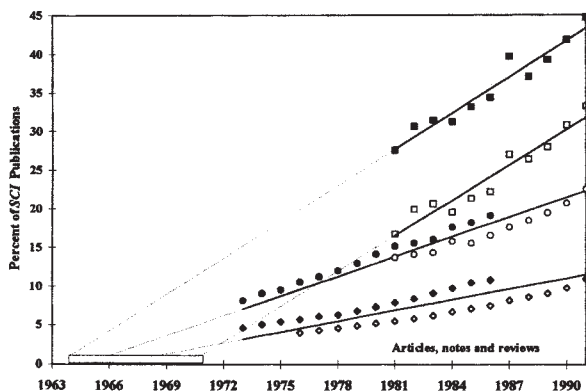
Survey teams are investigating sites in Argentina, South Africa, Australia, Spain, Russia, and the United States. Cronin expects four to eight major nations to join the project, and says a further meeting is planned tentatively for the autumn. This is likely to take place at the Paris headquarters of the United Nations Education, Scientific and Cultural Organization (Unesco), and would be the first step in the effort to recruit potential partners. **Tony Reichhardt**

Questions of collaboration

SIR — Collaboration between researchers has been increasing for many years, but whether this is the direct result of policy initiatives or reflects a process intrinsic to the global scientific community is an open question^{1,2}. Bibliometric evidence derived from the *Science Citation Index (SCI)* and produced by the Science Policy Research Unit and by CHI Research Inc. suggests that the rapid growth in institutional collaboration may have started in the late 1960s or early 1970s and may be due to processes intrinsic to the science system.

The figure below plots two types of institutional collaboration: international and domestic. The lower two lines depict the 1973 to 1991 trends in international collaboration with UK and US partners expressed as a percentage of UK and US *SCI* publications respectively. Because the counting techniques differ, the data from two sources were combined and an average linear regression line was calculated. The upper two lines display trends in UK industry collaborations with all domestic institutions and with UK universities only. These are expressed as a percentage of UK industry *SCI* publications. The regression lines for both the international and domestic collaborations have been extended backwards in time to indicate where they cross the x-axis.

In 1993 the British government released a White Paper that placed strong emphasis on formal and informal links between the public and private sectors. Irrespective of this and other government policies it is obvious from the data that domestic collaborations with UK industry are substantial and have increased throughout the 1980s. It can also be seen that international collaboration in the UK scientific community is greater than for the United States and has been so for some time.



Percentage of *SCI* publications with institutional collaborations. UK industry domestic collaborations: ■, UK industry—any domestic institution (source SPRU); □, UK industry—UK university (source SPRU). UK international collaborations: ● (source CHI Research Inc.), ○ (source SPRU). USA international collaborations: ◆ (source CHI Research Inc.); ◇ (source NSF).

However, the percentage of collaboration with both countries appears to have increased at a fairly constant rate since 1973. During this time there were two Conservative and one Labour administration in the United Kingdom and four Republican and two Democratic administrations in the United States, yet there is no evidence that policy changes in either country influenced the rate of international collaboration.

For some time the effects of European Commission programmes that promote collaboration within the European Union (EU) have been of general interest to policy-makers^{3,4}. A bibliometric analysis of UK–USA and UK–EU collaborations shows that, in 1973, about 1 per cent of UK papers involved an European partner and this proportion grew to approximately 7.5 per cent by 1991. By comparison, in 1973 about 2 per cent of UK papers involved a US partner and this grew to 6.5 per cent by 1991. In 1982, for the first time, a greater percentage of UK papers involved a collaboration with an EU than with a US institution. However, both growth rates were remarkably constant over this period and, on closer examination, it was found that since 1973 UK–EU collaborations had a higher growth rate than UK–US collaborations. At least at the macro level, there is no evidence for the effect of commission policy. The UK science system appears to be moving along its own trajectory.

The backward extrapolation of the collaboration trends suggest that both domestic and international collaborations may have grown at fairly constant rate for 20 to 25 years. This could be due to many factors such as the rapid expansion in international travel and telecommunications systems that began in the 1960s, the global diffusion of scientific knowledge

and personnel, the increase in the number of centres of scientific excellence, a growth in interdisciplinarity and the complex nature of research or perhaps the self-organizing nature of the science system. No matter what the causal relationship, there is an important lesson in these observations. Each country's scientific community is embedded in a global research system. At the macro level it appears to move in step with the invisible international dynamics that shape it quite independently of national policy. The global science system has a large inertia and institutional infrastruc-

ture linked in a vast network — so vast in fact that it may be difficult to reshape without implementing draconian measures. In order for governments to understand their national research enterprise within the context of the global science community, they must adopt holistic and systemic investigative methods. Finally, any policy maker must conclude that scientific collaboration, including public-private sector collaboration, is intrinsic to the modern scientific culture. Collaboration is the rule not the exception.

J. Sylvan Katz

Diana Hicks

Science Policy Research Unit,
University of Sussex,
Brighton BN1 9RF, UK

1. Katz, J. S. & Martin, B. R. *Research Policy* (in the press).
2. Moed, H. F., de Bruin, R. E. & Straathof, A. *Report EUR 14582* (Office for Official Publications of the European Commission, Luxembourg, 1992).
3. Narin, F. & Whitlow, E. S. *Report EUR 12900* (Office for Official Publications of the European Commission, Luxembourg, 1990).
4. Moed, H. F., de Bruin, R. E., Nederhof, A. J., van Raan, A. F. J. & Tijssen, R. J. W. *Report EUR 14581* (Office for Official Publications of the European Commission, Luxembourg, 1992).

Reading slides

SIR — I am surprised that John Tiffany (*Nature* 374, 671; 1995) has not tried to assess the effect of the entoptic phenomenon on readability of slides. This is not a new phenomenon — it was first described by Sauvages in 1763 and confirmed by Steinbuch in 1813 and Vierordt in 1856. The effect of blue light upon the perception of this phenomenon was described separately by Rood and Reuben in 1860. I wonder with his experience in ophthalmology whether or not he has attempted mass screening of his lecture audiences for this effect upon their concentration and retentive powers?

Marcus A. Sleightholm

Seacroft Hospital,
York Road, Leeds LS14 6UH, UK

Libel on Internet

SIR — In connection with the news item "Universities in clash over Internet 'libel' allegations" (*Nature* 374, 297; 1995), there is as yet no innocent dissemination defence on the statute book in the United Kingdom, although the government does have legislation in draft form. There are good arguments in favour of the prevailing US approach to this issue, namely that the plaintiff should have to show that the secondary distributor (the host institution) was at fault in permitting the libel to be published. Adopting that tack would stop bulletin board defamation becoming too much of a lawyers' honeypot.

Nick Bralthwaite

Clifford Chance, 200 Aldersgate Street,
London EC1A 4JJ, UK

The s(p)elling of apo(p)tosis

SIR — The recent correspondence¹ about the correct pronunciation of apoptosis prompted me to consider why this subject is so appealing. Thus, although apoptosis is undoubtedly of considerable importance, there is an increasing tendency to claim that it is involved in many cases of cell death where the evidence for its actual involvement is minimal. For example, the protein encoded by one of the two potential candidate genes claimed by different groups to be mutated in spinal muscular atrophy was recently named 'apoptosis inhibitory protein' on the basis of a minimal homology to baculoviral apoptosis inhibitory proteins and the claim that the morphology of dying cells in the disease was consistent with apoptosis². Similarly, it is often claimed (especially in grant requests) that Alzheimer's disease involves apoptotic death, although in fact it now seems more likely that such death is necrotic³.

Undoubtedly a major initial appeal of the apoptotic concept was the notion that it could be inhibited with protein synthesis inhibitors, leading to the exciting concept of an active suicide pathway requiring *de novo* protein synthesis (see for example ref. 4). Yet it now appears that in most cases the death pathway is constitutively primed and needs to be actively inhibited⁵. Thus a combination of a protein kinase inhibitor (staurosporine) and a protein synthesis inhibitor (cycloheximide) produces apoptotic death in most cell types. Had this experiment been performed in the early days of the apoptotic concept, it might have been regarded as akin to the statement that a person dies if shot in the head and heart simultaneously!

There is, of course, no doubt that apoptosis is vital in many systems. Yet our enthusiasm for this concept should not make us believe that it can explain all death phenomena particularly in human disease. In all cases, clear evidence will be required that apoptosis is actually occurring. We should concentrate on acquiring this evidence rather than worrying about how the term itself should be pronounced.

David S. Latchman

University College London Medical School,
Department of Molecular Pathology,
46 Cleveland Street, London W1P 6DB, UK

1. Finder, J. *Nature* **373**, 379 (1995).
2. Roy, N. et al. *Cell* **80**, 167–178 (1995).
3. Lassman, H. et al. *Acta neuropath.* **89**, 35–41 (1995).
4. Martin, D. P. et al. *J. Cell Biol.* **106**, 829–844 (1988).
5. Raff, M. C. et al. *Nature* **356**, 397–400 (1992).

SIR — I suggest that you ask the Greeks about the proper pronunciation of the word apoptosis. The word *απόπτωση* is still in use in the modern Greek language and its second p is very much pronounced just like the p in other Greek words

adopted by Western languages, such as helicopter, *Archaeopteryx*, optic, apterous, aptotic, symptomatic and so on.

J. G. Georgatsos

Laboratory of Biochemistry,
University of Thessaloniki, Greece

SIR — For good measure, helicopter is *herikoputa* in Japanese.

G. De Mey

Afrikalaan 11,
B-3080 Tervuren, Belgium

Necessary brain

SIR — John Godfrey's well-reasoned Commentary¹ neglected to elaborate the argument relating to the central role of the brain in judging the ontogeny of a human². That neglect provided the openings for the barrage of opposing statements supporting the position of the Roman Catholic church on the time of appearance of a human being^{3–7}.

One must distinguish between "human life" and the "life of a human person". Without a functioning brain there is no human person. Indeed, when the brain dies the person is regarded, legally and morally, as dead, even though the rest of this unique human body may still be alive and healthy. In fact, the heart, liver, kidneys and so on may legally be removed from this living body for transplantation into another. Clearly, a uniquely individual human body without a functional brain is regarded as human living tissues but not as a person. I believe even the Catholic church agrees with this view and policy.

At conception, regardless of the precise definition employed, there is no brain at all. When does the brain appear and develop to a stage one may regard as having the characteristics or qualities of a human person? The achievement of such characteristics is not a sudden, all-or-none phenomenon and may be regarded as developing in progressively meaningful stages. But there is surely a minimum stage below which the person could not begin to appear. This minimum stage would require not merely the appearance and organization of the neurons, especially in the forebrain, but also the development of synaptic and functional interconnections, to a level at which the brain exhibits physiological arousal, EEG responsiveness to a sensory input. This state is not achieved until well into the second trimester of fetal life and perhaps even later.

Before the achievement of a brain with at least the minimum characteristics for human personhood, the fetus is best regarded as living and developing human

tissue, no more a person than is an adult human individual whose brain is dead. The argument that there is a human person present before the minimum stage of brain (or even at conception), would have to be based on a religious or philosophical belief that the embryo or fetus is already a person even when there is no suitable brain.

Benjamin Libet

Department of Physiology,
University of California, San Francisco,
San Francisco,
California 94143-0444, USA

1. Godfrey, J. *Nature* **373**, 100 (1995).
2. Libet, B. *Science* **213**, 213 (1981).
3. Jones, D. A. & Haynie, D. T. *Nature* **373**, 652 (1995).
4. Herranz, G. *Nature* **373**, 652 (1995).
5. Jarmulowicz, M. *Nature* **373**, 652 (1995).
6. Johnson, W. V. *Nature* **374**, 10 (1995).
7. Montuenga, L. M. *Nature* **374**, 10 (1995).

Authors' rights

SIR — I have been asked to contribute a chapter to a book, and I am being asked to sign a contract with the publisher. It is a fairly standard contract for this sort of task: I agree to provide a chapter to their specifications, with illustrations, by a date they decide; I agree to obtain written permission to reproduce figures and tables from other published works; I agree to review copy-edited text and to review and correct proofs promptly (but not further revise them). Furthermore, I am asked to assign to the publisher all copyright in the material, both text and figures. (Even though I may spend hours drawing these figures on my personal computer, I shall have to ask permission from the publisher if I want to use them in another article.)

What do I get in return for all this? Almost nothing financial. (In this particular case, 5 per cent of receipts will be divided between about 60 authors.) Again, in this particular case I am only offered one copy of the book between myself and as many co-authors as I decide to co-opt. What does the publisher get? The text of a book that will no doubt bring in a nice sum of money over the years. What assurance do I get even that my article will be published? None. I am sure other authors will have shared my experience of learning some years after sweating over a chapter that the publisher has decided not to go ahead with the project.

Why do scientists put up with this? Does it reflect the low self-esteem in which we hold ourselves? Do we see ourselves as dogsbodies, working only for other people's gain? I have had enough — I may write the chapter, but I am not going to sign this contract. I hope other scientists will join me in beginning to protest about our exploitation by scientific publishers.

Kelth Frayn

Radcliffe Infirmary,
Oxford OX2 6HE, UK

Towards the synthetic dragonfly

The ability to predict and control the ways in which molecules interact and assemble through non-covalent interactions is now yielding a rich harvest in disciplines ranging from materials science to drug design.

It was not long after *Nature* published the structure of DNA that one began hearing less and less about biochemistry and more about its precocious offspring molecular biology. That it was 'chemistry' that was lost in this transition is revealing: the detailed chemical constitution of biomolecules, and the way in which this determines their function, became of less concern (because it was less accessible) than the function itself. Of course, high-resolution crystal structures are now changing that.

It is gratifying, then, to see to what extent chemical principles can now inform the science of molecular systems that, when not in fact derived from biology, are of comparable complexity. At *Nature's* conference last month*, chemistry was shown to have a 'systems-level' identity that previously only molecular biology could boast — but with the additional attraction that this is accessible to rational design, not a given consequence of natural selection.

If one accepts the dictionary definition of a system as 'a set of devices functioning together', it is soon clear why complex molecular systems cannot rely on the covalent bond alone. Weaker, more labile interactions become indispensable once the function of the system becomes at all sophisticated. One key reason for this was clear from several talks: whether the aim is to make new materials, new drugs or molecular memory devices, error correction is critical.

A covalent network might have the robustness of diamond, but mistakes in connectivity are frozen in. A complex system of any sort gains tremendously from the capacity for error correction. In this, as in so much else, biology provides the 'existence proof'. It means that systems-scale chemistry is the chemistry of the hydrogen bond, of metal-ligand interactions, of hydrophobic forces and steric effects. Organic synthesis will not quickly progress beyond the scale of palytoxin while it relies on forging covalent bonds; but once these weaker, annealable interactions are used, the scale of construction mushrooms.

Scale, however necessary for materials science, does not dictate function. To obtain functional molecular systems, the assembly process must be amenable to

control. But as Jean-Marie Lehn (Univ. Louis Pasteur) pointed out, this need not take us back to the square-one of laborious part-by-part assembly involving endless protecting and de-protecting steps, if instead we aspire towards 'instructed chemistry'.

This means that the parts contain the information necessary for their own assembly. It requires specificity of interaction, in some of Lehn's systems arising from preferred coordination geometries of metal ions and the ability of multidentate ligands to equilibrate — an error-correcting process — so as not to leave loose ends dangling. These most familiar of chemical principles alone are sufficient to allow an assembly of up to 19 molecular components to organize itself spontaneously in solution, within what Lehn calls an 'instructed mixture'. When the molecular building blocks are pre-programmed in this way, purity is no longer an issue, as extraneous molecules are simply ignored.

Lehn has been using these principles of molecular recognition and self-assembly to make molecular assemblies (organo-metallic complexes in the old order) based on linear arrays: ladders, stacks and grids. Fraser Stoddart (Univ. Birmingham) and Jean-Pierre Sauvage (Univ. Louis Pasteur) have focused on threading rather than stacking, which takes us into topology and knot theory. Even a unimolecular system is reasonably regarded as 'complex' when, as Sauvage described, it has been twisted into a trefoil knot. Threaded systems like the linked rings of catenanes retain a great deal of mobility between their components without falling apart, allowing them to acquire functionalities such as switching — all the more useful when accessible to external control via photochemistry or redox reactions.

In solution, this kind of coordination and host-guest chemistry is mediated by the solvent. In the solid state, design considerations are complicated by the exigencies of packing effects, which do not favour open, porous networks. Attempts by Jeff Moore (Univ. Illinois) to construct the hinged network of the thorium disilicide structure (which has unusual mechanical properties) using silver-ion coordination chemistry produced six such networks interwoven to fill up space.

Packing constraints in the solid state are central to the formation of the stacked pi-orbital systems that give rise to electrical conductivity in molecular conductors.

But as Paul Chaikin (Princeton Univ.) discussed, one-dimensional conductivity is just one of the extraordinary range of electronic and magnetic behaviours that these compounds can show. The tendency is for one-dimensional molecular conductors to undergo the Peierls instability at low temperatures, leading to dimerization along the stacks, an energy gap at the Fermi surface, and thus an insulating state. Enhancing electronic interactions between adjacent stacks, for example by applying pressure to reduce their separation, can suppress this instability and yield molecular superconductors. The charge-transfer salt (TMTSF)₂PF₆ exhibits not only this but also just about every other kind of electronic behaviour known in solid-state physics, including the quantum Hall effect seen otherwise only in strictly two-dimensional thin films.

Josef Michl (Univ. Colorado) made a persuasive case for why we should not ditch covalent chemistry too soon. He has collected an impressive construction kit of molecular components, including rigid rods, vertex connectors and support attachments, with which his aim is to build covalently linked scaffoldings as a sequence of stacked two-dimensional nets. The trick will be to balance stability (strong links) against reversibility (the error-correction imperative again); Michl's preferred solution at present is to explore the use of bonds between the organic components and mercury, which have a degree of covalency while retaining some lability.

Michl's strategy is essentially a thin-film technology, and that is one with which much experience has been gained already. Pre-organization of the building blocks would be effected by a process that is really nothing more than chemisorption of an ordered, incommensurate molecular layer, which is also what enables George Whitesides (Harvard Univ.) and others to make self-assembled monolayers. These organic films, in which linear molecules are attached at one end to an inorganic surface by a functional group with a strong affinity for that surface (thiols on gold, for example), have a high degree of order both in terms of packing and molecular alignment, providing extremely homogeneous surface properties. Intricate patterning of the films on a submicrometre scale is possible by a simple printing process, using an elastomeric stamp prepared from a mould fabricated by conventional lithography.

*The Science of Complex Molecular Systems, Paris, France, 27–28 April 1995.

Because a coating of a self-assembled monolayer conveys resistance to chemical etching, this process forms the basis of a microfabrication technology that has no need of electron beams and can generate features only 0.2 micrometres wide. It can, moreover, be applied to curved surfaces (which are simply rolled over the stamp), allowing circuit patterns to be imprinted on the surface of optical fibres — a reversal of the usual approach to integrated optoelectronics, in which the light must be brought to the chip.

Using cell machinery

There is much talk in micro- and nanofabrication circles of top-down and bottom-up approaches. In molecular science, the equivalent might be seen as the chemistry-up and biology-down strategies. All of the above begin armed with chemistry not so different from that outlined in Pauling's *The Chemical Bond*; but one might expect to get further (albeit at the cost of sacrificing some finesse) by hijacking the machinery of the cell. This, indeed, is the way in which one obtains antibody proteins that perform catalytic functions for chemical synthesis, on the whole by letting a stable analogue of the relevant transition state masquerade as an antigen. Richard Lerner (Scripps Research Institute) showed how the flexibility of this approach can be widened by 'reactive immunization', in which the antibodies are raised in response to a chemical reaction rather than a single molecular template.

Nadrian Seeman (New York Univ.) described how the enzymes that cut and stitch DNA in the cell can act as engineers of synthetic DNA-based architectures of astonishing complexity. By careful design of oligonucleotides Seeman creates junctions between three, four or more double-helical strands; and he has used ligases to assemble these into closed polyhedral cages with the topology of a cube and a truncated octahedron (the structure of the sodalite cage of zeolites). Sequence-specific degradation of individual strands using restriction enzymes produces fragments that can be analysed to prove the topological connectivity. With this sophisticated technology at hand, DNA zeolites are not an absurd proposal.

Perhaps these restriction enzymes will one day be replaced by Peter Dervan's (California Inst. Technology) synthetic molecules, which can bind into the minor groove of DNA and cleave it with high sequence selectivity. The molecules are based on the natural product distamycin, a crescent-shaped oligopeptide. The revelation that *two* such peptides will fit together in the minor groove has allowed Dervan to improve the selectivity by linking two distamycin derivatives via a hairpin turn. The hurdles that remain before molecules such as this perform 'gene surgery' *in vivo* will keep Dervan and his

competitors busy for many years yet.

Of course, there is nothing particularly new about conscripting enzymes for technological ends — the first biosensor, which used glucose oxidase to monitor blood sugar, was demonstrated in the 1960s. Today the same function can be performed by a device the size and shape of a pen (Anthony Turner, Cranfield Univ.). Biosensors continue to face challenges that could usefully occupy anyone working on complex molecular systems, who might be persuaded to provide enzyme mimics for recognition, self-assembly processes for fabrication, and electrically or optically responsive molecules for signal transduction.

But of all the technological ends that complex biomolecules currently serve, one of the most impressive is surely in optical memory devices. Robert Birge (Syracuse Univ.) described volumetric memories formed from stacks of bacteriorhodopsin, the light-harvesting membrane protein of *Halobacterium halobium*. Information is read in and out optically and is stored either layer by layer in 'page' mode or holographically, yielding associative memories more like those of neural-based systems. Site-directed mutagenesis can produce new excited states of the protein with lifetimes more conducive to energy storage than to the original function of transduction.

Light harvesting is one of nature's fortes, it seems, the ingenious solutions to which come no more spectacular than that of the 'storage ring' of photosynthetic bacteria (*Nature* **374**, 517; 1995). In attempting to design artificial systems that use light to drive chemical processes other than photosynthesis, we need not, according to Allen Bard (Univ. Texas), mimic such systems too literally. Rather, so-called integrated chemical systems for photoelectrochemical energy conversion can afford considerable latitude in the specifics, so long as the components perform the same basic processes of efficient light capture and long-lived charge separation. The latter is the most problematic. Photoinduced electron transfer from an excited chromophore is not difficult to achieve, but back-transfer takes place readily too. The ideal is to keep the excited donor and the acceptor far apart. In the chloroplast this is achieved by efficient electron relays; in the artificial systems that Bard described, all manner of tricks are employed. In one the photoactive component, a zinc porphyrin complex, is placed at the mouth of a zeolite channel (which it cannot penetrate), with the acceptor (a cluster of platinum metal) inside and safely out of reach: viologen ions act as a relay to shuttle electrons down the channel.

Even non-specific interactions such as those that govern colloid science can mediate notable feats of self-assembly and

self-organization (Jacob Israelachvili, Univ. California at Santa Barbara). The properties of micelles, bilayers and vesicles are now understood to an extent that allows their application in fields as diverse as crystal growth and drug delivery; and there is evidence that even so ornate an edifice as the cubic phase may play a role in cell biology. But injecting molecular recognition into such systems has tremendous potential benefits — for example allowing the construction of new cellular materials by the docking of streptavidin and biotin at the surfaces of liposomes (Israelachvili), or of colour-change biosensors by placing receptors for viruses at the surface of piezochromic polymerized vesicles (Helmut Ringsdorf, Johannes-Gutenberg Univ.).

Stanislav Leibler (Princeton Univ.) showed that, at the price of a few reasonable assumptions, as refractory a process as the self-assembly of microtubules from tubulin monomers during mitosis can be represented by a physical model that reproduces the microtubules' growth and contraction in their search for chromosomes. At this level, Leibler suggests, function can be retained without the need for a precise chemical blueprint for self-assembly.

Life is elsewhere

That complex molecular systems are now blurring the distinction between the living and the artificial has been one of the most widely trumpeted of chemistry's recent achievements. The forced evolution of polynucleotides promises to give us the key to the RNA world, as well as new drugs along the way. And exact replication of information-containing molecules without the assistance of enzymes has long been a minimal Holy Grail of those seeking to demonstrate the feasibility, if not necessarily the precise mechanism, of the chemical origin of life. By using synthetic nucleic-acid analogues, for example containing pyranosyl units in place of ribose, Albert Eschenmoser (ETH, Zürich) is beginning to answer questions about the uniqueness of DNA and RNA as the basis of life.

It is unlikely that the challenge of synthesizing a dragonfly would have been received without derision at *any* scientific conference a decade ago. But after what had gone before, George Whitesides was able to pose this problem in his conference summary without provoking too much raising of eyebrows. It is not so much that the objective is a realistic one, but that we can now see how to think about the problem. Artificial molecular motors, photoreceptors, metabolic systems: these are becoming realistic goals not so much because we know how to build versions modelled on biology but because we are starting to appreciate what we can afford to leave out. **Philip Ball**

Dendrites solidified by computer

Michel Rappaz and Wilfried Kurz

EVERY metallurgist (and almost every skier) is familiar with the appearance of dendritic crystals. Worldwide, as many as 10 billion metallic dendrites are produced in industry every second during solidification. Simulating every detail of the process in the computer has proved difficult, but two researchers from the US National Institute of Standards and Technology (NIST) have now developed a technique¹ that makes it possible to do so.

The omnipresent dendritic structure does not appear only in metals; it is also typical of snowflakes and of other materials such as ceramics, and is of course a key element in the neural system of animals and man (the human brain has about a hundred billion neurons, each being connected to neighbouring cells by dendrites and an axon). The name comes from the highly branched morphology: the Greek for tree is 'δένδρον' (dendron).

Transformations in inorganic materials being much simpler than in living matter, physical and mathematical modelling of crystal formation has made much progress in the past few decades. The key incentive is the prediction of properties. It is well

known to alpine skiers that the mechanical behaviour of a layer of snow, and thus the avalanche tendency, depends on the structure and interaction of the tiny ice dendrites. Similarly, the structure and interaction of metal dendrites control the behaviour of metallic alloys. A good example is provided by single crystalline turbine blades which are used in the harsh environment of jet engines at temperatures of some 1,000 °C and at up to 9,000 r.p.m. The high performance of these components is improved if the blades solidify with their dendrites all oriented along the same direction.

ally becomes unstable and breaks up into a dendritic morphology. The typical length scale of a dendrite is dictated by two antagonist phenomena. On one hand, diffusion (of heat and/or of solute elements) is most efficient when the interface becomes highly pointed, for the same reason that the elongated hull of a racing boat has a lower drag coefficient than a broad barge. On the other hand, curvature and the associated surface energy tend to make the crystal as blunt as possible. The snowflake in Fig. 1 is a typical result of such a compromise, the beauty being that its morphology also reflects the crystallographic nature of the molecular arrangement (hexagonal in this case).

For most metallic alloys, several phases can precipitate in between the arms of the dendrites depending on the alloy composition (an example is silicon platelets in an aluminium-silicon alloy for a car engine block). This mixture of ductile dendrites and hard interdendritic particles occurs at the length scale of the dendritic pattern and, to a large extent, conditions the ultimate mechanical behaviour of the part.

For many years, scientists struggled to understand the mechanisms of dendrite growth. Metallurgists of the eighteenth century were limited to drawing dendrites seen with the naked eye. Investigations carried out on transparent materials^{2,3} in combination with theoretical considerations^{4,5} helped to clarify the situation. But it was not until the pioneering work of Langer and Müller-Krumbhaar that a quantitative description was obtained for the growth of the dendrite tip⁶.

Although the dendrite tip plays a key role in determining the final microstructure formation, it is only a small part of a much larger structure. Other theories had to be used, each to describe the formation of other phenomena in the same dendrite: side branches, their evolution (coarsening), the segregation of solute elements and so forth.

With the advent of powerful computers and workstations, new routes have become available for the simulation of dendrite formation. James Warren and Bill Boettinger, respectively a physicist and a material scientist working at NIST, have now derived an elegant, comprehensive numerical technique for solving this problem¹. Rather as Cahn did in 1960 for spinodal decomposition⁷, they took into account the diffuse nature of the solid-liquid interface of metals. The transition from liquid to solid occurs over several atomic distances. Accordingly, they assumed the change between liquid and solid to be gradual over several mesh

points of a grid. In addition to a solute concentration, each point of the grid is characterized by a volume fraction of the liquid phase. This field, which gives the technique its name ('phase field'), is equal to unity if the grid point is fully liquid and zero after it has solidified. The solute diffusion equation is coupled to a phase-field

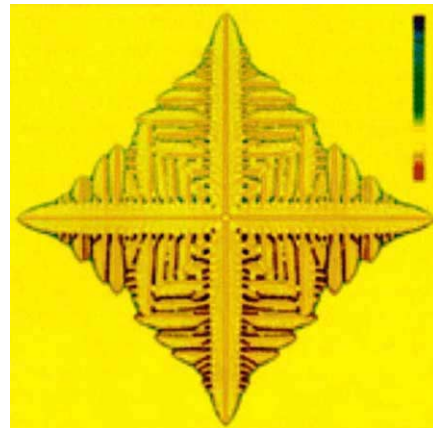


FIG. 2 Results of numerical modelling¹ of dendritic growth of an undercooled metal with cubic crystal structure. Colour bar shows concentration gradient. (Reprinted with permission from ref. 1.)

equation governing the evolution of the curvature of the system.

The image shown in Fig. 2 is the result of such a computation for a nickel-copper alloy using a grid of 750 by 750 points. The cubic structure of the solid metal is clearly reflected by the orientations of the dendrite branches, which grow at 90°. All the features observed experimentally for dendrites of alloys are present: dendrite tip growth, formation of secondary arms nearly perpendicular to the primary trunks, coarsening or coalescence of dendritic features, and microsegregation. So it seems that, although the grid size used in these computations is still much larger than the atomic scale, it successfully reproduces the behaviour of the solid-liquid interface of metals. The simulation, which is as yet limited to two-dimensional growth situations, definitely gives a new insight into the mechanisms of dendrite growth and is a substantial contribution to the problem of pattern formation. □

Michel Rappaz and Wilfried Kurz are at the Ecole Polytechnique Fédérale de Lausanne (EPFL), Laboratoire de Métallurgie Physique, MX-G, CH-1015 Lausanne, Switzerland.

1. Warren, J. A. & Boettinger, W. J. *Acta metall. Mater.* **43**, 689-703 (1995).
2. Papapetrou, A. Z. *Kristal.* **92**, 89 (1935).
3. Glicksman, M. E., Schaefer, R. J. & Ayers, J. D. *Metall. Trans.* **A7**, 1747 (1976).
4. Zener, C. *Trans. Metall. Soc. AIME* **167**, 550 (1946).
5. Ivantsov, G. P. *Dokl. Akad. Nauk SSSR* **58**, 567 (1947).
6. Langer, J. S. & Müller-Krumbhaar, H. *J. Cryst. Growth* **42**, 11 (1977).
7. Cahn, J. W. *Acta metall.* **8**, 554 (1960).

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

C Nuridsany & M. Perennou/SPL

FIG. 1 The familiar hexagonal, dendritic form of a snowflake.

known to alpine skiers that the mechanical behaviour of a layer of snow, and thus the avalanche tendency, depends on the structure and interaction of the tiny ice dendrites. Similarly, the structure and interaction of metal dendrites control the behaviour of metallic alloys. A good example is provided by single crystalline turbine blades which are used in the harsh environment of jet engines at temperatures of some 1,000 °C and at up to 9,000 r.p.m. The high performance of these components is improved if the blades solidify with their dendrites all oriented along the same direction.

The transformation of a liquid into solid has to start from a nucleus. Once a nucleus has formed, it starts to grow. But the interface between solid and liquid usu-

Serial engagement proposed

Mark M. Davis

THE sensitivity of an antigen-specific response by T cells is remarkable. Yet the number of cognate ligands formed by a peptide antigen complexed with a molecule of the major histocompatibility complex (MHC) on the antigen-presenting cell can be very small; and the interaction with the appropriate T-cell receptor (TCR) appears to be one of low affinity and short duration. How can this combination of high sensitivity and low affinity be reconciled?

On page 148 of this issue¹, Valitutti *et al.* present striking results which might offer a resolution to this puzzle. They show that TCR complexes disappear from the surface in numbers that are large multiples of the number of cognate peptide/MHC ligands available to them, and that irrelevant TCRs are largely unaffected. The model they propose is one of serial engagement, with the high 'off-rate' of the interaction allowing the same peptide/MHC combination to trigger up to 200 TCRs and thus amplify and sustain the activation response. The response, which is measured in this instance by production of the cytokine γ -interferon, correlates in magnitude and timing with the observed downregulation of the TCR. Thus the T cell seems to be 'counting' the number of receptors that have bound an appropriate ligand, formulating a proportional response, and then taking them out of service.

Off-rates

Serial engagement of a given peptide/MHC complex with what is typically a large excess of TCRs seems very likely, as the half-lives of the TCR/ligand interactions are of the order of 2–30 seconds (even when their affinities vary greatly)^{2–4}, and so over the 2–5 hour course of these cell-activation experiments one would expect that hundreds or even thousands of contacts would be made. These fast off-rates are not unique to TCR interactions, but instead seem to be a general property of all cell-surface receptors which 'receive' other cell-surface molecules⁵. Because of the highly polyvalent nature of molecular interactions between cells, the fast off-rates of these other molecules are probably important in allowing the cells to separate from each other. In the case of TCRs, however, whose contribution to the energy of the cell–cell conjugate seems negligible⁶, this property seems to have been put to another purpose.

Although previous work has shown that both TCRs and surface immunoglobulins largely disappear from the cell surface upon antibody cross-linking or (in the case

of immunoglobulins) ligand encounter, these events have been associated with wholesale 'capping' and internalization, and not with the specific, sequential process observed by Valitutti *et al.* Instead this seems more reminiscent of the receptor-mediated endocytosis described for epidermal growth factor and several other tyrosine-kinase receptors for soluble ligands⁷. The TCR is related to the tyrosine-kinase family of receptors by virtue of the tyrosine phosphorylation of its associated CD3 molecules which accompanies activation⁸.

Although many of the tyrosine-kinase receptors depend on soluble-ligand-induced dimerization for activation⁷, Valitutti *et al.* prefer a model of monovalent serial engagement for the TCR, and they do not see a role for receptor clustering. Nonetheless, it is clear that multimerization of TCR heterodimers and their non-covalently associated CD3 polypeptides can trigger T cells. The experiments showing this include cross-linking with antibodies against TCR or CD3 determinants and, more recently, cross-linking of chimaeric molecules consisting of a non-TCR extracellular domain linked to a CD3- ζ cytoplasmic dimer⁸. In addition, work with TCRs specific for fluorescence⁹ or arsonate¹⁰ has shown that these haptens can trigger specific T cells when highly multimerized but not in a monomeric or less multimeric state.

All of these data indicate that some multimerization involving CD3 domains is likely to be a component of activation, although reconciling a requirement for multimerization with the observed sensitivity of the T-cell response to its physiological ligand is not straightforward. I would argue that it is feasible, however, and that the data of Valitutti *et al.* do not preclude the involvement of dimerization in each serial activation and downregulation event.

To incorporate multimerization, one has to consider the question of how two TCRs can be aggregated by as few as 100 (refs 11–13) peptide/MHC ligands diluted by a sea of 50,000 or so complexes that are irrelevant. This is an especially acute problem if homogeneous peptide/MHC dimers are required to facilitate TCR dimerization on the T cell, as Brown *et al.*¹⁴ have proposed (based, in part, on their finding that MHC class II molecules form dimers upon crystallization). Such dimers of heterodimers have also been observed in immunoprecipitates of class II MHC molecules¹⁵. The problem is that the random juxtaposition of peptide/MHC complexes into dimers would mean that only 1 in 250,000 pairs (1/500)² would

have the same peptide, fewer than the number of MHC molecules on a given cell.

Enrichment

As Kupfer and Singer¹⁶ have pointed out, however, the juxtaposition of two cells would cause all of the complementary molecules on their respective cell surfaces to accumulate at the interface, or 'co-cap'. In a T-cell context, this would enrich significantly for the 'correct' peptide/MHC but not for irrelevant ones⁶. If the area in contact represents about 10 per cent of the available surface there could be a 10-fold enrichment; a peptide/MHC ligand present as 1 in 500 would therefore become 1 in 50 and the chances of having a homogeneous dimer would rise to 1 in 2,500. So there would be an average of one or a few homogeneous MHC/peptide dimers at the cell–cell interface, and by serial engagement this might be able to trigger tyrosine phosphorylation and the internalization of TCRs that the authors see. (Although it would require off-rates generally to be at the lower end of the spectrum, a reasonable possibility under physiological conditions.)

All in all, one can argue over whether the TCRs are disposed of one by one, or as cross-linked groups of two or more. Nevertheless the data of Valitutti *et al.* clearly reveal a new aspect of T-cell activation in which the same peptide/MHC complexes seem to be used over and over again to engage and trigger a much larger number of TCR molecules. How these observations are eventually integrated into a complete molecular and cellular picture of T-cell activation will be a story worth following. □

Mark M. Davis is in the Howard Hughes Medical Institute and the Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, California 94305-5428, USA.

- Valitutti, S., Müller, S., Cella, M., Padovan, E. & Lanzavecchia, A. *Nature* **375**, 148–151 (1995).
- Corr, M. *et al.* *Science* **265**, 946–949 (1994).
- Sykulev, Y. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **91**, 11487–11491 (1994).
- Matsui, K., Boniface, J. J., Steffner, P., Reay, P. A. & Davis, M. M. *Proc. natn. Acad. Sci. U.S.A.* **91**, 12862–12866 (1994).
- van der Merwe, P. & Barclay, A. N. *Trends biochem. Sci.* **19**, 354–358 (1994).
- Matsui, K. *et al.* *Science* **254**, 1788–1791 (1991).
- Fantl, W. J., Johnson, D. E. & Williams, L. T. A. *Rev. Biochem.* **62**, 453–481 (1993).
- Weiss, A. & Littman, D. R. *Cell* **76**, 263–274 (1994).
- Symer, D. E., Dintzis, R. Z., Diamond, D. J. & Dintzis, H. M. *J. exp. Med.* **176**, 1421–1430 (1992).
- Rao, A., Ko, W. W.-P., Faas, S. J. & Cantor, H. *Cell* **36**, 879–888 (1984).
- Demotz, S., Grey, H. M. & Sette, A. *Science* **249**, 1028–1031 (1990).
- Harding, C. V. & Unanue, E. R. *Nature* **346**, 574–577 (1990).
- Christinck, E. R., Luscher, M. A., Barber, B. H. & Williams, D. B. *Nature* **352**, 67–70 (1991).
- Brown, J. H. *et al.* *Nature* **364**, 33–39 (1993).
- Schafer, P. H. & Pierce, S. K. *Immunity* **2**, 699–708 (1994).
- Kupfer, A. & Singer, S. J. A. *Rev. Immun.* **7**, 309–337 (1989).

Dichotomous regulators

Stefan G. E. Roberts and Michael R. Green

THE transcription rate of a protein-coding gene is controlled by regulatory proteins that bind to specific DNA elements associated with the gene's promoter. In general, a particular regulator is an activator (which stimulates transcription) or a repressor (which abates it)¹⁻³. But several transcriptional regulators activate in one circumstance and repress in another. How? Something of the molecular mechanisms involved emerges on page 162 of this issue⁴, where Sauer *et al.* present evidence of the way in which the *Drosophila* protein Krüppel acts as such a dual-function regulator.

Promoter-specific transcriptional regulators must act through RNA polymerase II and general transcription factors that are required for the expression of virtually all genes¹. These general transcription factors, designated TFIIA to TFIIF, work by forming a huge preinitiation complex on the promoter. Some activators are known to stimulate transcription by increasing assembly of the preinitiation complex^{1,3}, but in principle activators could also act at subsequent steps in the transcription process (for example, by enhancing initiation or elongation of RNA polymerase II).

The *Drosophila* transcription factor Krüppel is encoded by a gap gene and is required for proper segmentation of the embryo. Krüppel converts from an activator to a repressor in a concentration-dependent manner^{5,6}; at low concentrations it is a monomer that activates transcription, whereas at higher concentrations it is a homodimer that binds to the same DNA sequence as the monomer but is a potent transcriptional repressor. To understand the molecular basis of this activator-repressor switch Sauer *et al.* analysed Krüppel's ability to engage in protein-protein interactions, and they now report that the two forms of Krüppel contact distinct general transcription factors.

The authors demonstrate, first, that the Krüppel monomer (the activator) specifically interacts with TFIIB. This comes as no surprise, in that TFIIB is known to have a central role in transcription activation. For example, several classes of activators, in particular those with acidic activation domains, interact directly and specifically with TFIIB both *in vitro*⁷⁻¹⁰ and *in vivo*¹¹. Mutagenic studies have provided evidence that, at least for some activators, the interaction with TFIIB is required for transcriptional activation⁸. And other studies have demonstrated that an activator can increase assembly of the preinitiation complex and hence

transcription^{7,9}. So the results of Sauer *et al.* are consistent with a large body of work on activation mechanisms.

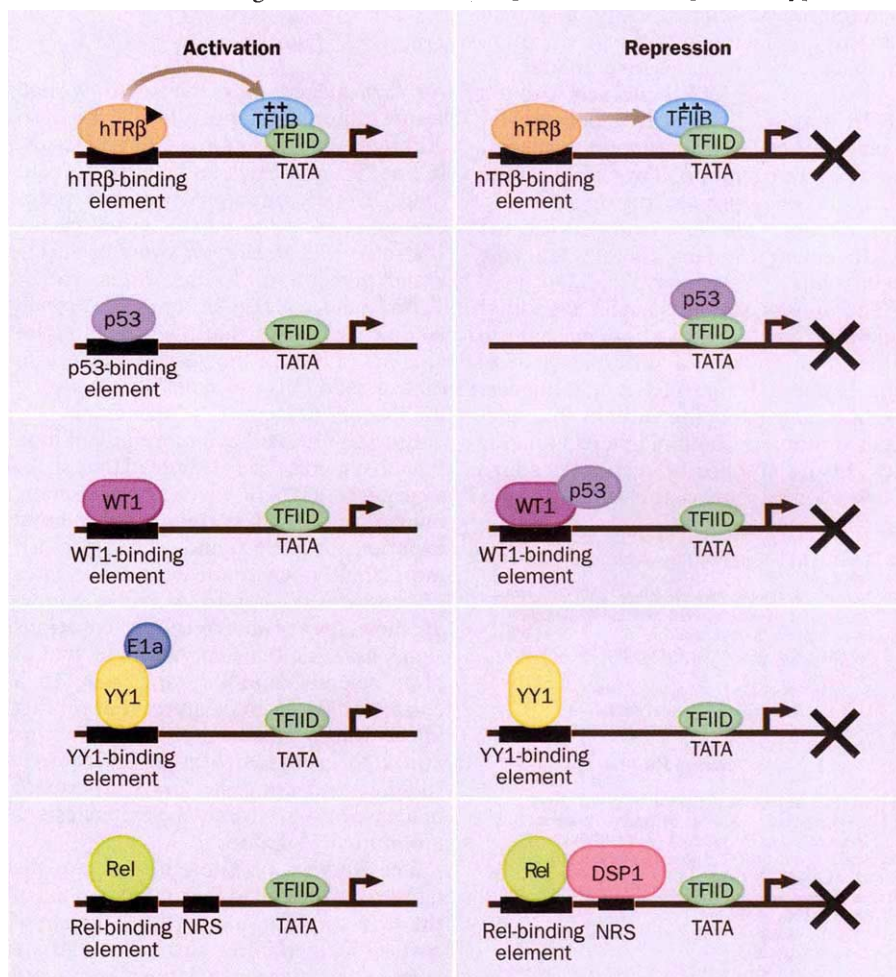
But if Krüppel can interact with TFIIB, why do Krüppel dimers repress rather than activate transcription? Sauer *et al.* show, second, that whereas the Krüppel monomers can bind TFIIB, the dimers do not. Instead, the dimer engages in a different protein-protein interaction, with the β -subunit of TFIIE (TFIIE β), a polypeptide with which Krüppel monomers cannot interact.

Thus Sauer and colleagues' main conclusion is that the monomer-dimer transition alters Krüppel's ability to interact with the general transcription machinery. Although the molecular details of this switch require further investigation, several possibilities seem likely. For example, Krüppel dimerization presumably masks the TFIIB binding surface and must

also create a TFIIE β binding site. This binding site could result from a dimer-induced conformational change or it could be formed by a contribution from each of the two Krüppel subunits.

The previous studies on activator mechanism provide a framework for understanding how the interaction between a Krüppel monomer and TFIIB leads to activation. But it is less obvious how interaction between the Krüppel dimer and TFIIE β represses transcription, primarily because this is the first report of an interaction between a transcriptional regulatory protein and this particular general transcription factor.

What do we know about TFIIE? TFIIE enters the preinitiation complex after RNA polymerase II and supports the subsequent assembly of TFIIF¹. TFIIF contains a kinase activity that phosphorylates the carboxy-terminal domain of the largest RNA polymerase II subunit, an event associated with transcription initiation¹. It is therefore possible that, rather than interfering with formation of the preinitiation complex, Krüppel dimers



Mechanisms of activator-repressor switches. Each of the transcriptional regulators is shown in either an activation (left) or repression (right) context. Switching can be caused by a ligand, thyroid hormone (black triangle) in the case of human thyroid hormone receptor- β (hTR β); by promoter architecture (p53); or by auxiliary factors (WT1, YY1 and Rel). Contacts with general transcription factors involved in the activation-repression switch are indicated. Only the relevant parts of the transcription complex are shown.

repress transcription by preventing RNA polymerase II from initiating. TFIIE is not necessary for transcription of some promoters¹², and it will be interesting to see if Krüppel can repress transcription in such cases, or whether it acts only at TFIIE-dependent promoters.

Several other transcriptional regulators can switch between activator and repressor, and the molecular basis for these transitions is diverse (see figure). One other well studied example also involves an interaction between the regulatory protein and TFIIB¹⁰. The human thyroid hormone receptor- β (hTR β) is converted from a repressor to an activator upon binding of its ligand, thyroid hormone. Both forms of hTR β bind to TFIIB, but the activator and repressor target distinct TFIIB regions. As an activator, the transcriptional activation domain of hTR β interacts with the so-called 'basic' region of TFIIB, the same part of TFIIB that is contacted by acidic activators⁸. As a repressor, the transcriptional repression domain of hTR β interacts with the TFIIB amino terminus, a portion of TFIIB that is required for subsequent entry of general transcription factors into the preinitiation complex¹³. So an attractive model for repression is that hTR β interacts with the TFIIB amino terminus to block further assembly of the preinitiation complex, and thus transcription. Thyroid hormone binds to the repressor region of hTR β and disrupts the interaction with the TFIIB amino terminus, thereby relieving repression.

The tumour suppressor p53 is another dual-function regulator whose mechanism appears to involve a different general transcription factor. p53 is a sequence-specific, DNA-binding protein that activates promoters containing a p53 binding site. In the absence of such a site, high levels of p53 repress transcription by a

mechanism that probably involves a non-productive interaction with TFIID¹⁴.

Other dual-function regulators are induced to switch activity by auxiliary factors that are recruited to the promoter either by a protein-protein or a protein-DNA interaction, and then act as coactivators or corepressors. For example, the product of the Wilms' tumour gene, WT1, is an activator in the absence of p53 but a repressor in its presence¹⁵. YY1, a mammalian sequence-specific DNA-binding protein with some homology to Krüppel, has complex transcriptional properties¹⁶ — when directed to the initiator region of the promoter, it can function as an activator; in contrast, when it operates from an upstream binding site, YY1 represses transcription, but it is converted back to an activator following binding of the adenovirus E1a protein. Finally, activators of the Rel family can be

converted to repressors by a cofactor designated DSP1, which binds to DNA sequences associated with Rel-binding sites in several cellular promoters¹⁷.

These examples underline the point that the activity of a transcriptional regulatory protein can depend upon both the promoter and the cellular context. How these different contexts lead to opposing effects on transcription is a challenging problem. However, the new work of Sauer *et al.* provides us with both a model and an experimental strategy for tackling other complex transcriptional regulatory systems. □

Stefan G. E. Roberts and Michael R. Green are in the HHMI Research Laboratories, Program in Molecular Medicine, University of Massachusetts Medical Center, 373 Plantation Street, Worcester, Massachusetts 01605, USA.

CONSCIOUSNESS

Knowing how, knowing where

Antonio R. Damasio

AT first glance only the word 'visual' seems to join two reports in this issue, so different are their approaches and styles. In one, Francis Crick and Christof Koch¹ (page 121) present a provocative hypothesis concerning the part of visual cortex whose activity we are *not* aware of during visual perception. In the other, Roger Tootell and colleagues² (page 139) provide new evidence that the human visual area V5 (a part of the visual cortices also known as MT) is especially active during the illusory perception of movement.

But the reports have more in common than first meets the eye. Indeed they share a purpose, that of discovering where in the human brain the correlates of mental experience can be found — or, to put it more bluntly, where the neural correlates of consciousness are to be found. Knowing how the brain engenders conscious states has long been the ultimate goal of brain science, and knowing how, to a considerable extent, requires that we first know where. Both groups of authors attack the problem bravely and quietly (neither ever using the terms experience or consciousness), and their conclusions are mutually supportive.

The findings of Tootell *et al.* are the latest contribution to the understanding of the extrastriate visual cortices, the set of cortices located in the surround of striate cortex (the latter is also known as primary visual cortex, or V1 or area 17). The functional mapping of these cortices, which began just over two decades ago with the work of Semir Zeki, has not ceased to reveal how different anatomical subsectors, identifiable on the basis of

their interconnectivity, are differentially active in response to varied properties of a visual stimulus, for instance colour or movement. Area V5 (or MT) in the monkey responds especially well to moving stimuli (although it also responds to the specific direction of movement), and the presumed human homologue of V5 has already been shown to respond to motion³.

The exciting news in Tootell and colleagues' report, however, is that activity in human V5 is also correlated with the illusion of motion. The investigators led their subjects into producing a visual motion after-effect known as the waterfall illusion. As a result, the subjects thought they saw motion although in reality they were being confronted by stationary stimuli; and while they had the illusory experience of visual motion, their presumed V5 areas were especially active. These fascinating results were obtained with one of the modern neuroimaging tools, functional magnetic resonance imaging (fMRI). Just as is the case with positron emission tomography (PET), fMRI provides an index of the activity in circumscribed neuron populations, estimated on the basis of local haemodynamic changes.

Crick and Koch's hypothesis addresses the intriguing question they pose in their title: "Are we aware of neural activity in primary visual cortex?". To understand

1. Zavel, L. & Reinberg, D. A. *Rev. Biochem.* **64**, 533–561 (1995).
2. Ptashne, M. *A Genetic Switch: Phage Lambda and Higher Organisms* 2nd edn (Cell Press/Blackwell Scientific, Cambridge, MA/Oxford, 1992).
3. Kingston, R. E. & Green, M. R. *Curr. Biol.* **4**, 325–332 (1994).
4. Sauer, F., Fondelli, J. D., Ohkuma, Y., Roeder, R. G. & Jackle, H. *Nature* **375**, 162–164 (1995).
5. Licht, J. D., Grosse, M. J., Figge, J. & Hansen, U. M. *Nature* **346**, 76–79 (1990).
6. Sauer, F. & Jackle, H. *Nature* **364**, 454–457 (1993).
7. Lin, Y.-S. & Green, M. R. *Cell* **64**, 971–981 (1991).
8. Roberts, S. G. E., Ha, I., Maldonado, E., Reinberg, D. & Green, M. R. *Nature* **363**, 741–744 (1993).
9. Kim, T. K. & Roeder, R. G. *Proc. natn. Acad. Sci. U.S.A.* **91**, 4170–4174 (1994).
10. Banihmad, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **90**, 8832–8836 (1993).
11. Colgan, J., Ashali, H. & Manley, J. L. *Molec. cell. Biol.* **15**, 2311–2320 (1995).
12. Parvin, J. D., Timmers, H. T. M. & Sharp, P. A. *Cell* **68**, 1135–1144 (1992).
13. Ha, I. *et al. Genes Dev.* **7**, 1021–1032 (1993).
14. Seto, E. *et al. Proc. natn. Acad. Sci. U.S.A.* **89**, 12028–12032 (1992).
15. Maheswaran, S. *et al. Proc. natn. Acad. Sci. U.S.A.* **90**, 5100–5104 (1993).
16. Shi, Y., Seto, E., Chang, L.-S. & Shenk, T. *Cell* **67**, 377–388 (1991).
17. Lehming, N. *et al. Nature* **371**, 175–179 (1994).

1. Crick, F. & Koch, C. *Nature* **375**, 121–123 (1995).
2. Tootell, R. B. H. *et al. Nature* **375**, 139–141 (1995).
3. Zeki, S. *et al. J. Neurosci.* **11**, 641–649 (1991).
4. Damasio, A. R. *Descartes' Error: Emotion, Reason and the Human Brain* (Putnam, New York, 1994).

their point, I found it useful to paraphrase the question in the following terms: "When we are conscious of what we see, do any of the neural correlates of the images we experience correspond to activity in V1?" Crick and Koch's answer is, none at all. V1 may be active, but the neural correlates for the image of which we are aware cannot be found in V1.

This anti-V1 manifesto is based on two lines of reasoning. The first is the result of reflecting on the connectational patterns of the visual cortices (gleaned from neuroanatomical studies in the monkey), and of considering the possible evolutionary role of awareness. Crick and Koch argue that the best possible use of the images that enter our awareness is in planning behaviour. Because planning depends on frontal cortices, then the contents of visual awareness must be made available to them by the appropriate connections; and because V5 or V4 have such connections, but V1 does not, V5 and V4 are in and V1 is out. The second line of reasoning draws on evidence from both psychophysical and physiological studies. For instance, we cannot turn off the colour constancy effect, something we might be able to do if we were aware of the vagaries of colour processing in V1; nor can we tell which part of a visual image comes from one eye or the other, although V1 is busy processing signals from each eye separately.

As Crick and Koch themselves acknowledge, their hypothesis is both subtle and difficult to test. But it is plausible and it may be a helpful guide, especially if they succeed in specifying precisely which visual areas contain correlates of visual awareness. If we grant that we are unaware of the results of whatever activity is going on in V1, are we then only aware of images based on the activity of 'fifth-tier' areas such as V4 and V5? Or can we be aware of ensemble activity in all early visual cortices minus V1? Might not the ensemble activity of these areas, which are rich in hierarchical and heterarchical projections among themselves, be a better candidate to support images of which we are aware than the outer extrastriate areas alone?

Taken together, the two reports help narrow down the search for the neural correlates of visual experience — that of Tootell and colleagues by demonstrating the involvement of an area in, of all things, a visual illusion, and that of Crick and Koch by proposing the exclusion of a possible area. Neither report claims, as well they should not, that the proposed areas are both necessary and sufficient for visual experience. The finding and the proposal before us specifically concern the neural correlates of the visual images in our minds.

But there is more to experience than that. When the subjects of Tootell *et al.*

perceived illusory movement, and knew, automatically and unshakably, that this was their own individual experience, many brain regions outside the visual cortices must have been necessarily co-active to support such subjectivity. A comprehensive mapping of the neural correlates of visual experience, at the level of large-scale systems, must then invoke non-visual brain regions. My own ex-

pectation is that the latter includes the entire set of structures, from the brain stem on up to the cortex, which map the body of the organism engaged in seeing and being aware of what is being seen⁴. □

Antonio R. Damasio is in the Department of Neurology, University of Iowa College of Medicine, Iowa City, Iowa 52242-1053, USA.

TECHNOLOGY

Superconducting superwires

Paul M. Grant

"ELECTRICITY is carried by wires. If it weren't for wires, electricity would be useless." I remember vividly it was with these words that my father, an avid Edisonophile and amateur radio experimenter, began my introduction as a young boy to the wonders of electrical science. It is easy to forget the central role this prosaic component plays in all aspects and uses of electricity; arguably, no other technology is so vital to the electric utility industry and its customers. Although copper and aluminium have met most of our wire needs for decades, the demand for conservation and more efficient use of electricity has brought renewed focus on superconductors — materials that have been with us for over 80 years, materials that could yield us veritable 'superwires'.

Despite their marvellous properties, superconductors still do not play much part in our everyday lives. Except for a few niche applications, notably electromagnets for NMR medical imaging, the attendant cost and inconvenience of helium cryogenic refrigeration have simply not justified their widespread use. The discovery in the 1980s of the cuprate perovskites, operable above liquid nitrogen temperature, raised great hopes that the economic balance would now shift in favour of broader application.

At a meeting last month*, progress towards this goal received a well publicized 'leg up' with the announcement of a world-record critical current density at 77 K (S. R. Foltyn, Los Alamos National Laboratory). This result was attained in 'thick' films of $\text{YBa}_2\text{Cu}_3\text{O}_{7-y}$ (Y-123) deposited on flexible metal tapes 5 cm long by 1 cm wide, an embodiment that could in principle be scaled up to kilometre lengths. ('Thick' is here used to differentiate from the 'thin' films of much less than a micrometre typically used in electronic applications.) These tapes, only a few tenths of a millimetre thick, were coated by pulsed laser deposition with Y-123 films of 1–2 μm .

The Y-123 films were capable of carrying more than 100 amperes of current with a critical current density in excess of one million A cm^{-2} , well above the requirements, for example, of electrical transmission lines. Of immense importance was that useful levels of critical current density were sustained at nearly 100,000 A cm^{-2} in a 5-tesla magnetic field, thus opening the way for practical magnet wire cooled by liquid nitrogen. The icing on the cake was the retention of this performance when the tapes were coiled to diameters as small as 2 cm. Thus we are approaching the performance achieved in epitaxial films or single crystals, but in a configuration that could be used as a wire. The announcement brings to light an alternative approach to high-transition-temperature (high- T_c) wire that has been under quiet development and improvement in laboratories throughout Japan, the United States and Europe for several years.

This alternative has to be viewed against the background of the present commercially available high- T_c wire. High-temperature superconductors are inherently brittle ceramics, pseudo-orthorhombic in symmetry and polycrystalline, properties not very friendly to wire formation. In a key 1988 paper¹, IBM workers showed that existence of low-angle grain boundaries between neighbouring microcrystals (low angle with respect to the copper-oxygen planes) was essential to obtaining maximal critical current, especially in magnetic field. Prospects for wire appeared dim.

But nature then granted two unexpected favours. The first was that cuprates could exist in thermodynamic equilibrium with metallic silver, permitting its use both as a matrix and as a sheath with which to dress the superconducting ceramic and provide the ductility necessary for wire formation (the wrapping on this first gift was the ease of diffusion of oxygen through silver, allowing post-processing of the cuprate in its final wire form). The second was the discovery of $\text{Bi}_2\text{Sr}_2\text{Ca}_2\text{Cu}_3\text{O}_{10}$ (Bi-2223) with a T_c of

*Spring Meeting of the Materials Research Society, San Francisco, California, USA, 19 April 1995.

110 K, and double bismuth oxide layers that had the micaceous property of shearing like a deck of playing cards into grain-aligned colonies on crushing and extrusion by the wire drawing and rolling process.

These benefices form the basis for the oxygen (postprocessing)-powder (Pb-stabilized Bi-2223)-in-tube (Ag), or OPIT, wire technology now being commercialized by Sumitomo, American Superconductor and Intermagnetics General. The final product is a silver-clad tape half a millimetre in thickness, about half a centimetre in width, and available in lengths now exceeding one kilometre (ref. 2). Critical current densities in zero magnetic field are length-dependent, ranging from 20,000 to 30,000 A cm⁻² at 77 K for a few metres, to about 12,000 at the kilometre scale. However, with respect to the proverbial third wish — in this case, robustness in even modest magnetic field — nature was more capricious. The same double bismuth oxide layer so helpful in producing grain alignment serves to degrade critical current to uselessly low values in magnetic fields of only a few tenths of a tesla at 77 K.

Y-123 has long been the material of choice for liquid nitrogen operation because it does retain high critical currents under substantial magnetic fields in epitaxial film or single crystal form. This property is a direct consequence of the stronger vertical coupling between copper oxygen planes provided by the CuO chains in its crystal structure as opposed to (Pb,Bi)-2223. On the other hand, this same feature inhibits the self-grain alignment yielded serendipitously by (Pb,Bi)-2223 and has defeated all attempts to use the OPIT process to make wire out of it. The processability difference between the two materials is readily discernible when grinding them in a mortar and pestle: (Pb,Bi)-2223 gives a greasy, graphitic feel, whereas Y-123 is akin to sand. The goal has been to find a process whereby long lengths of a suitable flexible substrate can be coated with grain-aligned Y-123.

Early on, efforts were made to deposit thick films of Y-123 on nickel alloy tapes of special composition (Hastelloy) chosen to match the thermal expansion coefficient of Y-123. An immediate problem was the diffusion of nickel into the Y-123 film at high processing temperatures, solved by interposing a thin 'buffer layer' of a suitable oxide, most often yttria-stabilized cubic zirconia (YSZ), between the tape substrate and superconductor. Around 1991, a 1-m conductor (in utility industry parlance, a conductor is made out of wires, a cable is what houses the conductor) containing 100 tapes of this design was constructed by Sumitomo Electric, demonstrating that it could in principle be manufactured on large scales. But because the Y-123 layer was polycrys-

talline with a large distribution of wide-angle grain boundaries, the critical current performance both in and out of magnetic field was disappointing (K. Fujino, results presented at ISS 6 conference, Hiroshima, 1993).

In 1992, workers³ from Fujikura Ltd obtained a dramatic improvement in properties using ion-beam-assisted deposition (IBAD), a process pioneered by IBM⁴ before the advent of high-*T_c*, to texture the YSZ buffer layer during its deposition. The process can be thought of as preferential atomic sandblasting whereby those crystallographic directions not presenting clear 'channels' to the impinging ion beam (usually argon) are eroded away. The resulting oriented buffer layer then seeds the growth of a deposited Y-123 film in a quasi-epitaxial manner. On centimetre-square samples, the Fujikura group achieved a zero-field critical current at 77 K of 2.5×10^5 A cm⁻² and astonishing values only an order of magnitude lower at 5 T applied field, outperforming OPIT even at this early date.

Will the combination of IBAD and thick films render OPIT obsolete? There is no question that (Pb,Bi)-2223 will not be able to match its performance at 77 K. Nonetheless, a number of technical problems remain, not the least of which is the feasibility of scaling up a complex thin-film deposition process to industrial volumes and areas while retaining laboratory-level performance. The outlook is hopeful — there are many similarities to methods used today for the mass production of magnetic recording disks for the personal computer industry. Indeed, a preliminary study⁵ of manufacturing costs indicates no obvious technical stumbling blocks and suggests that if sufficient demand for high-performance superconducting wire at 77 K were engendered, the Y-123 tape could be produced and sold at a price competitive with current low-*T_c* wire products used in MRI machines.

So the future of IBAD/thick-film superconducting tape really involves an economic, rather than technical, risk, and we will have to wait and see which entrepreneurs or companies surface with the courage (or foolhardiness!) to take it on in expectation of the reward returned by a potentially large market. But isn't this what Adam Smith told us capitalism was supposed to be all about? □

Paul M. Grant is at the Electric Power Research Institute, 3412 Hillview Avenue, Palo Alto, California 94303, USA.

1. Dimos, D. *et al.* *Phys. Rev. Lett.* **61**, 218 (1988).
2. Malozemoff, A. *et al.* *Cryogenics* **32** (ICMC Suppl.), 478 (1992).
3. Iijima, Y. *et al.* *Appl. Phys. Lett.* **60**, 769 (1992).
4. See, L. *et al.* *Appl. Phys. Lett.* **47**, 932 (1985).
5. Hammond, R. H. *et al.* *Rep. 1995 DOE Wire Development Workshop*, St Petersburg, Florida, January 1995.

Twisted currents

ELECTRONS have spin. They can spin in either of two directions, and the two need not be equivalent. The electrons emitted by radioactive β -decay spin preferentially to the left, for example. Daedalus has uses for a beam of left-spinning electrons. In an electron microscope, it should be preferentially scattered by left-handed molecules; a neat way to spot chiral regions in a biological specimen. In a television cathode ray tube, it should emit left-circularly polarized light. But a β -emitter is not ideal as an electron gun for instrumentation or imaging purposes. So Daedalus proposes another source of polarized electrons.

It is based on the spin-flip laser. In a magnetic field, the conduction electrons in indium antimonide can spin either with the field or against it. The ones spinning against the field can shed their higher energy by executing a 'spin flip' which emits polarized infrared radiation and leaves them spinning with the field, like the others. Soon all the conduction electrons will be spinning the same way. Pass a current through the semiconductor, and you could flush them out. Spin being conserved, they would retain their spin throughout the rest of the circuit. The result: circularly polarized electricity.

The first large-scale applications of polarized electricity will be polarized LEDs and electroluminescent panels, the first efficient sources of polarized light. Many previously unfeasible optical schemes will become practical: glare-free screens and control panels, car headlamps which illuminate but can't dazzle, stereo television and so on.

But Daedalus has further uses up his sleeve. Chiral electricity should carry out chiral electrochemistry. At the moment this can be done after a fashion at electrodes with molecularly chiral surfaces; but polarized electricity should do it perfectly. Specific optical isomers will be accessible directly in 100 per cent yield. The pharmaceutical industry, always seeking to match its products to the chemistry of life, will be delighted. And with luck, a chiral storage battery will also work. It will accumulate polarized electricity, and release it on demand. The user will be able to buy the new product in battery form, rather than having to maintain a tricky spin-flip generator. The only snag is that polarized electricity will take time to get going. A voltage is established at the speed of light; but the electrons of the circuit amble along at a much slower pace. It may take hours to flush out a new circuit with polarized electrons, and reap their lopsided benefits.

David Jones

How long is a giant sperm?

SIR — In 1950, Cooper¹ reported that the spermatozoon of *Drosophila melanogaster* “proves to be a most impressive gamete”, having a length of 1.76 mm, and commented, “just what the selective value of such a tremendously elongated spermatozoon may be is a matter of pure conjecture.” These sperm were considered giants by any standard, being, for example, approximately 300 times longer than human spermatozoa. It has recently become clear, however, that by *Drosophila* standards, the sperm of *D. melanogaster* are tiny.

Here we report that males of *D. bifurca*, a distant relative of *D. melanogaster*, produce sperm that are 58.29 ± 0.66 mm long ($N=3$ males); each gamete is approximately 20 times longer than the flies manufacturing them. Sperm length for this species was unambiguously determined using a dissection technique described elsewhere². These sperm are two-and-a-half times longer than those of the sibling species, *D. hydei*, which held the previous record for sperm length³.

The existence of such giant sperm poses a conundrum for evolutionary biology, for it seems to contradict established views on the evolution of sexes (anisogamy) and of male sexual strategy. Whereas males are expected to produce vast numbers of low-investment gametes which they use with minimal frugality, males of some giant-sperm-producing *Drosophila* use their sperm with female-like judiciousness, carefully partitioning their limited sperm among successive females². Moreover, giant sperm are relatively costly to manufacture. First, comparative analyses of 11 *Drosophila* species reveal a trade-off between sperm length and the number of sperm produced and transferred to females by males (S. P., unpublished data). Second, to manufacture giant sperm, males must develop very long, energetically expensive testes. For example, the testes of *D. bifurca* are each 67 mm long and comprise nearly 11% of a male's total dry body mass (compare with 5% in *D. melanogaster*) (S. P., unpublished data). Third, studies of *D. pachea*⁴ (sperm length 16.53 ± 0.29 mm) and *D. hydei*³ (sperm length 23.32 ± 0.51 mm) have suggested that time required to grow their large testes is causally responsible for the unusually delayed rates of male sexual maturation observed in those species. We therefore examined sex-specific ages of

reproductive maturity in *D. bifurca* ($N=30$ per sex), and found that females became sexually mature, on average, within 7 days of the final moult (eclosion), while males required 17 days (see ref. 4 for methods), thereby supporting the hypothesis that producing giant sperm constrains age at first reproduction.

What, then, are the selective advantages of producing such long sperm? It is likely either that longer tails confer some advantage in the competition to fertilize ova² or that they may contribute to some post-fertilization function, as illustrated by Karr⁵. In a recent report, Bressac *et al.*⁶ claim giant *Drosophila* sperm represent “another way of being anisogamous”, in which each sperm has a high prospect of fusion with an egg and serves the post-fertilization function of provisioning the zygote. These authors report that the sperm tail remains substantially intact throughout embryonic development and is ultimately eliminated with metabolic wastes on hatching, and therefore suggest that the function of giant sperm is to ensure the input of paternal mitochondrial DNA.

Although intriguing, this claim is based on limited data — that all or most of the giant sperm tail enters the egg in three species: *D. melanogaster*⁵, *D. littoralis* and *D. hydei*, and that the efficiency with which sperm are used by females increases with increasing sperm length. Moreover, their claim appears unlikely for three reasons. First, detailed examination of fertilization in several *Drosophila* species with giant sperm reveals that, whereas the entire

sperm tail is incorporated into the egg in some species, only a small fragment enters in other species. For example, in *D. bifurca*, < 3 mm of the 58.3-mm-long sperm enters the egg (T. L. Karr and S. P., unpublished data). Any hypothesis addressing post-fertilization functions are therefore refuted, at least as a general explanation for sperm length evolution in *Drosophila*. Second, comparing among species, the number of progeny produced per copulation declines drastically with increasing sperm length (compare 300–700 progeny per mating in *D. melanogaster*⁷ versus about 70 in *D. hydei*⁶). Any increase in the efficiency with which females use longer sperm may ameliorate this reduction in fitness, but will be unlikely to compensate for it. Third, given the significant costs associated with the production of longer sperm, if the primary benefit is to increase the paternal representation of mitochondrial DNA in offspring, as Bressac *et al.*⁶ suggest, this would represent an unlikely instance of ‘selfish’ mitochondrial genes winning out over nuclear genes. Further studies are required to resolve this enigma. Although our perspective of what represents a giant sperm has increased substantially, its selective value is just as conjectural as it was 45 years ago.

Scott Pitnick

Department of Biological Sciences,
Bowling Green State University,
Bowling Green, Ohio 43403, USA

Greg S. Spicer

Institute of Molecular Medical Sciences,
Palo Alto, California 94306, USA

Therese A. Markow

Department of Zoology,
Arizona State University, Tempe,
Arizona 85287-1501, USA

A coral mitochondrial *mutS* gene

SIR — A high rate of evolution is a much mentioned characteristic of animal mitochondrial DNA (mtDNA). However, this claim is based on observations made principally for mammals, where mtDNA nucleotide substitutions occur at frequencies 5–17 times those of nuclear DNAs^{1–3}. The differences in nucleotide substitution rates in mammals have been attributed to a lack of mtDNA repair mechanisms¹: only enzymes that could carry out excision repair of DNA damaged by oxidation have been reported to occur in mammalian mitochondria⁴. Whether or not invertebrate mtDNAs evolve faster than nuclear DNAs is somewhat controversial^{2,5}. Therefore it is of considerable interest that in the mtDNA of the octocoral, *Sarcophyton glaucum* (phylum Cnidaria) we have identified a gene for a homologue of MutS, a component of the bacterial MutSLH mismatch repair pathway. The octocoral *mutS* homologue is also of note because it is the only protein gene other than the con-

stant set of 12 or 13 energy-pathway protein genes⁶ so far reported for metazoan mtDNAs.

The Gram-negative bacterial MutSLH mismatch repair pathway is primarily a back-up for correcting replication errors that evade proofreading, although it also acts on mismatched bases in recombination intermediates^{7,8}. As well as MutS, this

- Cooper, K. W. in *Biology of Drosophila* (ed. Demerec, M.) 1–61 (Wiley, New York, 1950).
- Pitnick, S. & Markow, T. A. *Am. Nat.* **143**, 785–819 (1994).
- Pitnick, S. & Markow, T. A. *Proc. natn. Acad. Sci. U.S.A.* **91**, 9277–9281 (1994).
- Pitnick, S. *Behav. Ecol. Sociobiol.* **33**, 383–391 (1993).
- Karr, T. L. *Mech. Dev.* **34**, 101–112 (1991).
- Bressac, C., Fleury, A. & Lachaise, D. *Proc. natn. Acad. Sci. U.S.A.* **91**, 10399–10402 (1994).
- Gilbert, D. G. *J. Insect Physiol.* **27**, 641–650 (1981).

- Brown, W.M. in *Evolution of Genes and Proteins* (eds Nei, M. & Koehn, R.K.) 62–88 (Sinauer Associates, Sunderland, Massachusetts, 1983).
- Moritz, C., Dawling, T. & Brown, W.M. *A. Rev. Ecol. Syst.* **18**, 269–292 (1987).
- Wallace, D.C. *et al. Curr. Genet.* **12**, 81–90 (1987).
- Tomkinson, A.E., Bonk, R.T., Kim, J., Bartfeld, N. & Lin, S. *Nucleic Acids Res.* **18**, 929–935 (1990).
- Monnerot, M., Solignac, M. & Wolstenholme, D.R. *J. molec. Evol.* **30**, 500–508 (1990).
- Wolstenholme, D.R. *Int. Rev. Cytol.* **141**, 173–216 (1992).
- Modrich, P. *A. Rev. Genet.* **25**, 229–253 (1991).
- Modrich, P. *Science* **266**, 1959–1960 (1994).
- Varlet, I., Pallard, C., Radman, M., Moreau, J. & de Wind, N. *Nucleic Acids Res.* **22**, 5723–5728 (1994).
- Chi, N.W. & Kolodner, R.D. *J. Biol. Chem.* **269**, 29984–29992 (1994).
- Ross-Macdonald, P. & Roeder, G.S. *Cell* **79**, 1069–1080 (1994).
- Clark-Walker, G.D. *Curr. Genet.* **20**, 195–198 (1991).

[illegible]

b

Homo sapiens Dup
Mus musculus Rep3
Schizosaccharomyces pombe SW14
Saccharomyces cerevisiae MSH3
Saccharomyces cerevisiae MSH4
Homo sapiens MSH2
Mus musculus MSH2
Xenopus laevis MSH2
Saccharomyces cerevisiae MSH2
Escherichia coli MutS
Salmonella typhimurium MutS
Azotobacter vinelandii MutS
Streptococcus pneumoniae HexA
Saccharomyces cerevisiae MSH1
Sarcophyton glaucum mtMSH

There is evidence that, in yeast, mitochondrial genes evolve more slowly than nuclear genes¹². It has been shown recently that the yeast nuclear-encoded MutS-related protein MSH1 is imported into mitochondria, and that this protein

preferentially binds to mismatch-containing DNA¹⁰. Further, mutations in the yeast *MSH1* gene resulted in mtDNA accumulating point mutations¹⁰. These observations support the view that *S.g.* mtMSH functions in mismatch repair, and suggest that it would be worthwhile to examine relative rates of nucleotide substitution in mitochondrial and nuclear genes of octocorals. It seems likely that a mismatch recognition function for *S.g.* mtMSH would be concerned only with post-replicative repair, as there is no evidence for intermolecular recombination in metazoan mtDNA².

Finally, it can be inferred from the size of the *S. glaucum* mitochondrial genome (18.2 kb) and the likely inclusion of all protein and rRNA genes common to other metazoan mtDNAs⁶, that additional components of a mitochondrial mismatch repair pathway in this organism would be nuclear-encoded.

Geneviève A. Pont-Kingdon

Norichika A. Okada

Jane L. Macfarlane

C. Timothy Beagley

David R. Wolstenholme*

Department of Biology,
University of Utah, Salt Lake City,
Utah 84112, USA

Thomas Cavalier-Smith

G. Desmond Clark-Walker

Research School of Biological Sciences,
Australian National University,
Canberra, ACT 2601, Australia

*To whom correspondence should be addressed.

Climate and CCN

SIR — Toumi *et al.*¹ estimate a substantial cooling effect on climate caused by an increase in the formation of aerosol particles and cloud condensation nuclei (CCN). This is associated with a possible increase in OH radicals in the troposphere, in turn caused by increased penetration of ultraviolet radiation through the depleted ozone layer.

The authors should be credited for pointing out a potentially important process. However, we feel that a quantitative estimate of such a complex series of events requires a more sophisticated analysis. In particular, we would like to point out the following, potentially important effects not considered by the authors. They would generally tend to counteract the estimated negative climate forcing.

(1) The depletion of stratospheric ozone is most pronounced in winter and early spring in mid and high latitudes, whereas the impact on climate of changes in the aerosol abundance is largest in summer and in low latitudes. Unfortunately, Toumi *et al.* give no information about the seasonal dependence of the effect.

(2) An increase in tropospheric OH would lead to a more than proportional increase in H₂O₂, leading to a more effi-

cient oxidation of SO₂ in cloud droplets. This would suppress the concentration of SO₂ outside the cloud and thereby tend to decrease the formation of new particles through the OH reaction. The authors argue, referring to our own previous work², that oxidation in cloud droplets is not limited by oxidants (like H₂O₂) but by the occurrence of clouds. However, we never made such a categorical claim. We believe that H₂O₂ may well be in short supply, especially during the darker months, as well as in regions affected by pollution.

(3) Even in remote regions, the net impact on the CCN population of changes in OH is not easily estimated. If the hypothesis by Raes *et al.*³ is correct, most CCN in the marine environment are formed in the free troposphere where, especially in the subtropics, OH may be the dominant pathway for SO₂ oxidation. If this is the case, the formation of H₂SO₄ (and CCN) would be limited by the vertical transport of dimethyl sulphide and the subsequent formation of SO₂, and not by the concentration of OH.

We would also like to point out an error in Fig. 1 of ref. 1. If the numbers presented were correct, the total rate of aerosol formation would be less than 0.01 Tg S per yr, far too small a figure to make this process at all significant.

Henning Rodhe

Department of Meteorology,
Stockholm University,
S-10691 Stockholm, Sweden

Paul Crutzen

Max-Planck-Institut für Chemie,
6500 Mainz, Germany

TOUMI ET AL. REPLY — We agree with the points raised by Rodhe and Crutzen. However, we still maintain that the main uncertainty is the relationship between H₂SO₄ and cloud droplets and not the extent to which H₂SO₄ production has changed. We are encouraged by a recent modelling study⁴, confirming that the number of cloud condensation nuclei may indeed be very sensitive to OH. We apologize for the error in Fig. 1 caption: it should read $\times 10^3 \text{ mol cm}^{-3} \text{ s}^{-1}$ (error in the proof).

Regarding the specific points raised by Rodhe and Crutzen:

(1) The relative importance of cloud scattering increases with solar zenith angle. This may offset the seasonal variation of the ozone trend. Note that other modifications of stratospheric ozone (for example, by the QBO, solar cycle, volcanic eruptions) have different seasonal and latitudinal dependencies. The main objective of our paper was to demonstrate the potential importance of the mechanism relative to the direct radiative effect from stratospheric O₃ changes.

(2) Our radiative forcing calculation was intended for marine stratus clouds.

We cannot quantify the effect of SO₂ oxidation limited by H₂O₂, but do not think that that this effect is large globally.

(3) The variation of SO₂ oxidation with cloud lifetime and OH as a function of altitude and latitude was included in the model.

Ralf Toumi

Department of Physics,
Imperial College,
London SW7 2BZ, UK

Slimane Bekki, Kathy Law

Centre for Atmospheric Science,
Department of Chemistry,
University of Cambridge,
Cambridge CB2 1EW, UK

1. Toumi, R. *et al.* *Nature* **372**, 348–351 (1994).
2. Langner, J., Rodhe, H., Crutzen, P. J. & Zimmermann, P. *Nature* **359**, 712–716 (1992).
3. Raes, F., Van Dingenen, R., Wilson, J. & Saltelli, A. in *DMS: Ocean, Atmosphere & Climate* (eds Restelli, G. & Angeletti, G.) 331–322 (Kluwer, Dordrecht, 1993).
4. Russell, L. M., Pandis, S. N. & Seinfeld, J. H. *J. geophys. Res.* **99**, 20989–21003 (1994).

Amazon molly and Muller's ratchet

SIR — The Amazon molly, *Poecilia formosa*, is a gynogenetic fish and a hybrid between *P. mexicana* and *P. latipinna*. The all-female *P. formosa* need to mate with males of *P. mexicana* or *P. latipinna*, as parthenogenetic egg development can start only after sperm fusion. It is generally believed that the sperm chromosomes are subsequently eliminated and that there is no genetic contribution from the male. Scharl *et al.*¹ now claim that heterospecific males may occasionally contribute genetically to the offspring of *P. formosa* females. They reason that such a process could counteract the operation of Muller's ratchet (accumulation of deleterious mutations in clonally reproducing organisms). We think, however, that such a conclusion is not supported by their data and unlikely for theoretical reasons.

To state that a species avoids Muller's ratchet, it is necessary to show that deleterious mutations are replaced by 'healthy' ones through recombination. However, Scharl *et al.* have presented no evidence for such a process. They found that the genetic trait "black pigmentation", which is typical for the black molly males used in routine laboratory crosses with *P. formosa*, is occasionally present in *P. formosa* offspring, and that such offspring contained heritable microchromosomes in addition to the normal chromosomes. These results do not indicate that (parts of) these supernumeraries recombine with the *P. formosa* standard chromosomes. Therefore, the title "incorporation of subgenomic amounts of DNA as compensation for mutational load in a gynogenetic fish" is not justified.

Even if one assumes that paternal genetic material becomes a stable in-

tegrate of the *P. formosa* genome, it is doubtful whether this phenomenon is strong enough to counteract Muller's ratchet. It seems likely that only one or a few parts (regions close to a centromere) of the black molly genome can be incorporated into heritable microchromosomes. Otherwise, one would expect high frequencies of different microchromosomes in the laboratory population, which is not the case. Therefore, the number of genes that can be revitalized seems limited.

Schartl *et al.* also argue that paternal leakage into an asexual strain offers a genetic advantage to the heterospecific males. However, this is only true when the asexual and sexual strains exchange genes in both directions. In this system, however, *P. formosa* females are an evolutionary dead-end: their genes are never transmitted back into the host species. As a result, males that mate with *P. formosa* cannot enjoy increased fitness in their

own species and therefore cannot be selectively rewarded for their genetic contribution to the asexuals.

We do not disagree with the possibility of occasional introgression of paternal characters into *P. formosa*: this has already been suggested by Haskins *et al.*². It would explain why *P. formosa*, as well as other gynogenetic fish^{3,4}, are more genetically diverse, as would be expected for clonal animals. Yet the data do not allow Schartl *et al.* to conclude that the Amazon molly can escape Muller's ratchet.

Leo W. Beukeboom, Rolf P. Weinzierl

Nico K. Michiels

Arbeitsgruppe Michiels, Max-Planck-Institut für Verhaltensphysiologie, D-82319 Seewiesen, Germany

1. Schartl, M. *et al.* *Nature* **373**, 68–71 (1995).
2. Haskins, C.P., Haskins, E.F. & Hewitt, R.E. *Evolution* **14**, 473–483 (1960).
3. Vrijenhoek, R.C. *Am. Zool.* **19**, 787–797 (1979).
4. Moore, W.S. in *Evolutionary Genetics of Fishes* (ed. Turner, B.J.) 329–398 (Plenum, New York, 1984).

Community response to IRONEX

SIR — The iron fertilization experiment in the sea near the Galapagos Islands in mid-October 1993 (IRONEX)¹ showed that the cell division rate of unenclosed phytoplankton, initially high, was quickly enhanced by Fe addition, as had been found in bottle experiments in other high nutrient (nitrate)-low chlorophyll regions. However, the increase of phytoplankton concentration during the days after the IRONEX experiment was not merely an issue of cell physiology², but was a system response resulting in a new balance between cell division rate and losses from grazing and physical factors. I suggest that even for the first 3 days after fertilization (days 299–301) the grazers continued to keep this offshore phytoplankton in check despite the increase in the rates of photosynthesis and cell division. In addition, the remineralization of carbon by the food web during this time helps to explain the small effect of the iron enrichment on the CO₂ system³.

The grazing loss, not measured during IRONEX, is calculated here for the upper ~20 m during the first 3 days after fertilization, using estimates of photosynthesis (from ¹⁴C uptake over 24 h starting within 1 h of local noon; Fig. 3 of ref. 1) and change of phytoplankton biomass (from microscopy: C values kindly supplied by K. Buck, K. J. Coale and S. J. Tanner). The near-constant concentrations of the inert tracer SF₆, added to the Fe-enriched water (column 6 of table; values kindly provided by A. J. Watson), show that losses of phytoplankton from physical dilution were negligible. Also, sinking of the dominant small cells¹ can be discounted.

Using the ¹⁴C uptake data during the 3 days as a measure of particulate

autotrophic production (83 µg C l⁻¹, sum of column 3) and subtracting the concentration increase (7 or 8 µg C l⁻¹, column 4), 75 µg C l⁻¹ of phytoplankton vanished (sum of column 5). Thus, nine-tenths of the newly formed material (75/83) appear to have been eaten. In contrast, ref. 1 estimated that only 43 µg C l⁻¹ was grazed by the microzooplankton over the first 9 days of the experiment. There will have been additional feeding by larger (meso-) zooplankton. Further, the grazers seem to have 'caught-up' with the enhanced primary production by day 301 (column 4) so that the community reached a new balance between phytoplankton cell division and mortality in 2–3 days.

The increase of phytoplankton biomass was largely due to very small cells¹ which are principally consumed by protozoans (microzooplankton). But the biomass of these grazers increased only from 4.7 to 7.3 µg C l⁻¹ over 9 days¹, mainly from days 298 to 301 (K. J. Coale, personal communication), despite their large food consumption. Apparently, the population

dynamics of the microzooplankton continued to be tightly controlled by their predators, and so the only lasting biomass increase from growth due to the fertilization was probably among the larger zooplankton. Such an increase, without a change in chlorophyll concentration, is clearly seen in the seasonal studies in the high nutrient-low chlorophyll subarctic Pacific, when during the late spring and summer the light-limitation of the phytoplankton is removed⁴. This biomass increase would have been difficult to measure during the IRONEX cruise.

Also, elsewhere in the open, undisturbed eastern equatorial Pacific, the phytoplankton population dynamics are understood to be controlled by grazing in an Fe-limited ecosystem^{5,6}; the grazing remineralizes iron, which is immediately taken up by the phytoplankton⁷. Obviously, a balance of rates of cell division and mortality (grazing) alone cannot maintain the observed temporal and regional stability of phytoplankton concentrations. Because the high cell division rates (estimated to have been 1.0 per day initially and 1.5 per day on day 301)¹ show that iron supply did not markedly constrain the phytoplankton physiology, stabilizing factor(s) such as feeding threshold and a low half-saturation constant of the grazing response⁸ must have led to a new equilibrium differing surprisingly little from that in the ambient waters.

In conclusion, considering the zooplankton-caused limits as the proximate control of phytoplankton concentrations during IRONEX is profoundly different from referring to phytoplankton physiology² or the carrying capacity of the system limited by settling of particulate Fe (ref. 1). The algal population dynamics at the IRONEX station on the Galapagos shelf¹, however, behaves differently from populations in the open sea. I believe one reason to be the Fe-induced larger size of the cells and the ensuing problems for the grazers, just as on any other shelf where there are phytoplankton blooms.

Karl Banse

School of Oceanography, University of Washington, Box 357940, Seattle, Washington 98195, USA

1. Martin, J. H. *et al.* *Nature* **371**, 123–129 (1994).
2. Kolber, Z. S. *et al.* *Nature* **371**, 145–149 (1994).
3. Watson, A. J. *et al.* *Nature* **371**, 143–145 (1994).
4. Frost, B.W. *Prog. Oceanogr.* **32**, 17–56 (1993).
5. Cullen, J. J. *et al.* *J. geophys. Res.* **97**, 639–654 (1992).
6. Price, N. M. *et al.* *Limnol. Oceanogr.* **39**, 520–534 (1994).
7. Hutchins, D. A. & Bruland, K. W. *Mar. Ecol. Prog. Ser.* **110**, 259–269 (1994).
8. Frost, B. W. & Franzen, N. C. *Mar. Ecol. Prog. Ser.* **83**, 291–303 (1992).

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. Priority will be given to letters of fewer than 500 words and five references.

Spectral analysis

Michael J. Morgan

Colour: Art and Science. By Trevor Lamb and Janine Bourriau. Cambridge University Press: 1995. Pp. 237. £35, \$59.95 (hbk); £17.95, \$24.95 (pbk).

SOME weeks ago a portrait of the English chemist John Dalton appeared on the cover of *Science*. In the same issue, an article described the genetic analysis of material from Dalton's preserved eye, and the discovery that he lacked the gene for the cone pigment sensitive to medium-wavelength light. The finding is in agreement with Dalton's own description of his anomalous colour vision. The science of colour has now advanced to the point where the genetic code for cone pigments can be related to psychological observations with a fair degree of confidence — a remarkable achievement, and one for anti-reductionists to contemplate gloomily.

Colour: Art and Science tells the story of colour wonderfully well. I cannot recommend this beautifully produced and eclectic book too highly. The list of contributors and their subject matter should be sufficient to lure the reader: David Bomford (history of art), Bridget Riley (colour for the painter), Malcolm Longair (physics of colour), Denis Baylor (physiology), John Mollon (evolution and psychophysics), Peter Parks (natural history of colour), John Gage (colour and culture) and John Lyons (colour in language). Colour is a subject that exercises the whole intellect and sensibility of a person. Most aspects of colour are touched on in this book, and the line-up of contributors could hardly be improved. An interesting omission is the professional philosopher. The influence of language and culture on colour perception is discussed in detail, but the old war-horse "how do I know that my colour is the same as yours?" is not seriously confronted, perhaps because it is difficult to say anything new about the topic. The dreaded Q word (*qualia*) does not make an appearance. Space should perhaps have been given to those such as Nicholas Humphrey who think that *qualia* are a key issue, and to those such as Daniel Dennett who think that they are not. The only other omission is of a good index. The one on offer is rather minimal: there is, for example, no reference to Semir Zeki, even though his work is described in the text.

By and large, readers are left to make connections between different chapters; this should present them with few difficulties, for the signposts are clear and the journey interesting. One connection that struck me is that between Mollon's account of the colour theory of Gaspard

Monge and Riley's fascinating account of Claude Monet's "*enveloppe*", which she traces to an innovation of Titan, "which put European painting on a road that no other tradition had trod before, or, for that matter, since". The innovation was to treat colour as a property of the whole scene rather than of isolated objects. Monet returned to the Creuse valley to complete a painting that he had started in the winter (or early spring) and found himself unable to continue because an oak tree had sprouted leaves and spoiled the '*enveloppe*'. He had the leaves removed.



Colour uncoordinated — artistic representations of rainbows are not always very accurate. This curious example, "*Fishing for Souls*" by Adriaen Pietersz van de Venne (1589–1662) shows yellow outside green, with no hint of red.

Exactly a hundred years before, in 1789, Monge, Monet's fellow countryman, had given what was in effect the first demonstration of colour constancy to the Academie Royale des Sciences. Monge established that the colour of an object was a property of other objects in the scene, indeed of the whole scene. Incidentally, I cannot agree with Mollon that Monge wrote "with the frightening clarity that can be achieved only in the French language". Has Mollon tried recently to read Maurice Merleau-Ponty or Michel Foucault? And was it not a Frenchman, Victor Hugo, who spoke of "*l'obscurité indispensable*"? French was the preferred language of diplomacy because it lends itself to high-sounding ambiguities. There is enough of Citizen Chauvelin in me to think that English is an acceptable means of clear expression, especially in the hands of clear thinkers such as Mollon and the other contributors to this book.

We understand in the main how the first stages of colour vision are accom-

plished by the selective absorption of different wavelengths of light by three classes of photoreceptor. But do we understand what happens before and after this event? What happens before is the province of optics, and here Longair assures us that all is well. To the non-physicist it is deeply mysterious that a photon can pass through two slits at once and then interfere with itself (as Paul Dirac put it) but apparently "[n]one of this causes physicists to lose any sleep at all". I suspect some irony here, or else I should be moved to complain about the Mogadon Theory of Light. The explanatory powers of modern physicists are enormously impressive, but surely there is no reason to think that they will be any less susceptible to improvement than the achievements of Newton and Maxwell?

What happens after the cones? Here we know much less. Even the way in which midgut ganglion cells achieve their colour

opponency is not entirely clear, and how the brain separates out the chromatic from the achromatic component of the signal is obscure. A colour-processing area of the visual cortex has been described by Zeki and his colleagues, and damage to this area perhaps produces the neurological condition of central achromatopsia, in which the patient can discriminate wavelengths but not name coloured surfaces. The influence of culture and language on colour naming remains a difficult topic; the most attractive theory — that every language chooses its colour names from the same gamut of 11 basic colours — is not established beyond doubt. One day there will be a grand unified theory that explains, among other things, why it is forbidden to call brown shoes '*brun*' in French.

Since there does not seem to have been a trouble-maker at the Darwin Conference on which the book is based, I shall dissent here from Baylor's statement that "the rich colours that we see are inven-

tions of the nervous system". It was not the brain that made the sky different from the grass, or the flushed face of excitement different from the pale face of anger. I believe that colours are concepts that bind together objects sharing common properties of which wavelength is a signature. If we were provided from birth with a set of spectacles that turned colours into their complementaries, I believe we would still see grass as green. Fortunately, this conjecture is extremely difficult to disprove. None of which causes me to lose any sleep at all. □

Michael J. Morgan is at the Institute of Ophthalmology and Department of Anatomy and Developmental Biology, University of London, London EC1V 9EL, UK.

Down to Earth

Alan Krupnick

Small is Stupid: Blowing the Whistle on the Greens. By Wilfred Beckerman. Duckworth: 1995. Pp. 202. £20.

MOST economists are born contrarians. We delight in the perverse outcome and the exception that proves the rule or, better, that points the way to a more general rule. Wilfred Beckerman is a contrarian's contrarian, writing not one but two, as he says, "politically incorrect" books on how the environmental movement has it wrong. The first, *In Defence of Economic Growth*, written in 1974, took on the anti-growth movement spawned by Donella Meadows *et al.*'s *The Limits to Growth*, a deeply flawed book forecasting a bleak future as a result of unchecked pollution, natural resource exhaustion and population growth. His second book now takes for its title a wordplay on "the great anti-growth guru" E. F. Schumacher's paean to environmentalism, *Small is Beautiful*. And here in the juxtaposition of the two titles lies the book's greatest contribution and its failings.

Common ground

Beckerman believes that his is one of the few voices of reason in the wilderness and mounts a deliberately provocative, strident attack against favourite positions of the 'greens', views widely accepted by a substantial fraction of the general public and government officials. Unfortunately, he displays so much relish (and skill) for the task that he misses a major opportunity to point out, and ultimately to seek, common ground with the environmental community. If social welfare is to improve, this common ground must be sought and held, even at the expense of a contrarian's delight.

As an economist, I wholeheartedly

agree with much of this book. I cheer when Beckerman takes on, in turn, the myth that we are running out of resources, those groups who would have developing countries sacrifice economic growth for the developed countries' ideas of environmental protection, those advancing the notion that sustainable development is really a new and better substitute for the economist's "tired refrain" of maximizing social welfare and those who seem to care more about intergenerational equity than improving equity — the distribution of wealth and welfare — in this generation.

Take his arguments about the trade-off between environment and growth in developing countries. He argues successfully that the main environmental problems in these countries are not those that occupy most environmental groups and governments in developed countries. In the United States, enormous sums of money are spent on the clean-up of hazardous-waste sites that pose barely any risk to people or the environment. Billions of dollars are also spent every year on further cutting air pollution and water pollution, although the latter at least is of immediate risk primarily to recreational fisheries. By contrast, developing countries face immediate, obvious health threats from diarrhoea and other diseases associated with poor sanitation (faced by at least two billion people), polluted drinking water (faced by one billion) and other social ills that are a result of poverty and can therefore be alleviated by economic growth.

His arguments against the notion that we are running out of nonrenewable resources are familiar, but cannot be stressed enough because this notion is so intuitive, yet empirically so wrong. It is true that rapid population and economic growth have resulted in enormous increases in the extraction of metals, oil, coal and other nonrenewable resources. It seems obvious that because there is, for all practical purposes, a fixed supply of such resources on Earth, sooner or later we will run out of them.

Yet what is meant here by a 'resource' is unclear. A material may have been recoverable 20 years ago only from deposits of high concentrations. But with technological progress itself driven by growing scarcity of deposits at these concentrations, the same material can now be extracted at far lower concentrations. In the absence of major discoveries (which is generally the case), how else can one explain the fact that, as Beckerman artfully shows, reserves between 1970 and 1989 have increased for many important minerals, even while cumulative consumption over this period has, in some cases, exceeded 1970 reserves? For oil the story is similar, with the depth and adversity of conditions under which extraction can profitably take place increasing over time.

With physical scarcity so hard to define, economists argue that the only reasonable measure is price: if supplies are getting harder to find or more costly to extract, prices will rise, all other things being equal. Economists find that the prices of these resources have been falling or are holding steady over time, in absolute terms and after correcting for all other influences on price. The implication is that we are not running out of resources.

Shakier ground

Beckerman is on shakier grounds with his characterization of the global-warming debate. I share some of his scepticism about the extent to which the indisputable increase in greenhouse gases will cause physical and economic hardship, particularly in the light of equilibrating mechanisms in the atmosphere and oceans and nature's incredible ability to adapt to changing circumstances.

It is when he mischaracterizes and then rails against the 'precautionary principle', which has guided the Clinton administration's policy toward global warming, that I part company. Beckerman sets up a straw man for this principle: take dramatic action now at very high costs as long as there is the remotest possibility of serious harm in the distant future. By contrast, the Clinton administration and most developed countries have taken the position that, given the chance of serious harm, some low-cost actions should be taken now, while continuing to fund studies of the threat as well as the research and development to mitigate it. This has involved setting relatively modest, cost-effective goals for reining in emissions of greenhouse gases and attempting to implement win-win policies (in other words, policies that on other grounds are sensible and improve social welfare). The reluctance of most countries participating in last month's climate conference in Berlin to embrace the timetables and targets advanced by several nations for the beginning of the next century, shows the limits to the precautions that governments and the public are willing to take.

Economists are fond of saying there are no free lunches. But this truism applies to argumentation as well. By focusing on the extreme positions of others, the centre may be missed. And it so happens that a lot is going on at the centre of the environmental movement. The most important trend is the growing understanding of and support for the use of market or economic incentive mechanisms in the conduct of environmental policy.

Environmental groups routinely refer to the 'hidden costs' of polluting activities on society and the need to 'internalize' them, which means, ultimately, uncovering and estimating these costs and using emissions fees or other instruments to price the polluting activity at its social

costs rather than just its private costs. A variety of studies, in both the United States and Europe, of the social costs of electricity and of transport, some by environmental groups, some by economists, attest to the convergence of views. The economist A. G. Pigou, who in 1932 first published the idea of taxing pollution, would surely be pleased.

In addition, in the United States, the idea of using tradeable pollution-permits as a way of meeting environmental goals cost-effectively is strongly supported by economists and many environmental groups. Using this approach, the government gives away or sells permits-to-pollute that can then be bought and sold among polluters, so those who face the highest cost of abatement can buy permits sold by those who find it cheap to abate, thereby keeping the costs of abating to a minimum, at least in theory. This idea has been turned into policies for reducing sulphur-dioxide emissions from power plants in the United States, as embodied in the

US Clean Air Act of 1990, and for meeting ambient ozone targets in Los Angeles.

Readers should not take the above examples as evidence of the end of environmental extremism. Far from it. Beckerman is right to "blow the whistle on the greens" for the lack of balance in their views, their failure to recognize the major environmental improvements of the past two decades, and the moral superiority of their tone. But however catchy *Small Is Stupid* is as a book title, the arrogance it implies can only make the common ground harder to find and hold. I urge readers to look beyond the rhetoric and read some excellent arguments for a less alarming future and for policies that promise to increase social welfare without sacrificing economic growth. □

Alan Krupnick, formerly a senior economist on the Council of Economic Advisors in the Clinton administration, is at Resources for the Future, 1616 P Street NW, Washington DC 20036, USA.

The truth in the balance

Julian Morris

Environmental Gore: A Constructive Response to Earth in the Balance. Edited by John Baden. *Pacific Research Institute*. Pp. 270. \$21.95, £20. Distributed in the UK by the Institute of Economic Affairs, 2 Lord North Street, London SW1P 3LB.

JUST a quick glance at the cover of this book, which depicts a fluorescent green Dali-esque molten flat Earth sliding over the edge of a precipice, and readers will realize they are in for a contrarian treat. In *Earth in the Balance*, US Vice-President Albert Gore presented the definitive 'environmentally correct' vision of the future, in which "[t]he task of saving the Earth's environment must and will become the central organizing principle of the post Cold-War world". In *Environmental Gore*, five economists, three climate scientists, two lawyers, an ecologist, a businessman, a philosopher, a judge and a (deceased) Democratic politician, give their "constructive response", identifying four main problems with Gore's analysis.

Gore misinterprets and distorts scientific evidence, exaggerating the importance of certain environmental problems. In particular, he focuses on global warming as the most important environmental problem now facing mankind. However, as the chapters by Robert Balling and Richard Lindzen point out, the evidence for human-induced global climate change is not that strong.

First, the 0.5 °C increase in global temperature observed over the past century is

less than that predicted by Gore and most of the general circulatory models on which he rests his case. That may be because until recently these models did not account for the cooling effect of aerosols emitted alongside greenhouse gases: the new model produced by the UK Meteorological Office, which takes these aerosols into account, shows a better fit with the data.

Second, the 'urban heat island' effect and 'desertification' cause an upward bias in the temperature data. Third, the temperature data correlate better with changes in solar activity than with the predictions of the general circulatory models.

Fourth, although most of the increase in atmospheric greenhouse-gas concentration has occurred since 1940, most (75 per cent) of the observed warming occurred before this date.

Fifth, although Gore claims that increases in planetary temperature accelerated during the 1980s (and even the new UK Meteorological Office model predicts such a change), microwave emissions from oxygen molecules in the lower atmosphere, recorded by NASA (National Aeronautics Space Administration) from its low-orbiting satellites, suggest only a very modest 0.001 °C rise in temperature between 1979 and 1990. Clearly, the data do not indicate that a climate catastrophe is on the way.

Gore apparently misunderstands how human societies work, and so fails to identify correctly the underlying causes of environmental problems. Gus diZerega offers a compelling account of human

societies as self-organizing complex systems. Institutions such as property rights, the market and the law, constantly evolving and adapting to satisfy human wants, provide a framework of rules within which individuals living in complex societies can engage in mutually advantageous exchanges. By contrast, Gore conceives of societies' institutions in purely instrumental fashion; they are the playthings of philosopher-kings, to be ordered according to a masterplan. The real environmental problems faced by people around the world have mostly been caused or exacerbated by attempts to impose order on institutions from above. In sub-Saharan Africa, land degradation is endemic because governments prevent individuals from owning land. In the western United States, bison were driven close to extinction because they could not be owned until they were shot. Individuals need incentives to invest in the protection of the natural environment, and the best incentive someone can have is the expectation that investments made today will bring rewards tomorrow (or next year, or in a century).

Gore claims that human beings must undergo "a change in our essential character" in order to save the environment, both for ourselves (and our children) and for the environment itself. To support this claim, he attempts to trace the history of mankind's irresponsible use of the natural environment. His prime suspect is Francis Bacon, the English philosopher. Yet, as James Lennox points out, it was Bacon who asserted that "Man is but the servant and interpreter of nature". Why Gore should vilify Bacon, seeing in him the root of the evil technophilic world of automation-consumers in which Gore claims we live, remains a mystery. This misunderstanding is unfortunate because it leads Gore to reject what Lennox calls "the philosophical tradition responsible for [our] unprecedented standard of living".

What is more, as Evaristo De Miranda shows, attempting to save the environment for its own sake may be counter-productive. People living in the Amazon rainforest have shaped their environment to suit their needs and, as a result, have increased the level of genetic diversity in the forest. Their system of shifting agriculture has resulted in local fluctuations in the height of the tree canopy and thereby enabled new species of flora and fauna to survive. The recent enclosure of some of these areas as national parks, to preserve the 'natural' environment, has not only increased population pressure in the Amazon (a problem aggravated by subsidized agriculture and logging concessions), but has also reduced the level of genetic diversity.

Although Gore is well aware of the influence of interest groups on the politi-

cal process, he seems to believe that only evil companies feeding the consumer-junkies of our depraved society are guilty of behaving in such a self-interested manner. The notion that perhaps environmental groups might be less than squeaky clean in this respect is omitted from his analysis. This is a serious and telling omission. Serious, because it is these groups (along with their partners in crime, big business) who have been dictating the discourse on environmental regulation for the past 25 years. Telling, because Gore is, of course, keen to keep these people sweet: they command votes, and what is a politician without votes?

The authors in this collection mostly hit the bull's-eye. Their task is, of course, made that much easier by the ineptness of much of Gore's analysis — Lennox declines even to comment on "Gore's embarrassing attempts to bolster his claims with references to split brain research, chaos theory, the Heisenberg uncertainty principle, the general theory of relativity, and so on [because] these references display the muddled and undigested lack of comprehension of a bad undergraduate essay and are not worth responding to". But I have a few minor quibbles. The chapter on ozone depletion is poorly written and badly laid out. Barrett Walker's arguments in his chapter on trade are muddled. (Among other

things, he seems to be in favour of including natural-resource depletion in national accounts, whereas Robert Hahn, in an earlier chapter, spends a good deal of time discussing the problems associated with resource accounting.) And many of the themes in the book are repeated (for example, the first four chapters all take issue with Gore's spiritualism). These problems aside, *Environmental Gore* is a good read, and a must for anyone who has read Gore's diabolical diatribe.

Gore claims that "the evidence of an ecological Kristallnacht is as clear as the sound of shattering glass in Berlin". Offended by this analogy, some critics have described Gore himself as an 'ecofascist'. Although this overstates the case, the authors of *Environmental Gore* show clearly that, if enacted, his policy of national environmental socialism would "inadvertently impoverish the future of the world for our children as well as for the rest of life". □

Julian Morris is at Clare College, Cambridge CB2 1TL, UK.

Seeding the Universe

Neil Turok

Cosmic Strings and Other Topological Defects. By A. Vilenkin and E. P. S. Shellard. Cambridge University Press: 1995. Pp. 459. £60, \$100.

THE central concepts in unified field theories of particle physics are symmetry and symmetry-breaking. These twin notions underlie the Salam-Weinberg theory of electroweak unification, which particle-collider experiments have recently confirmed with remarkable precision. They also form the basis of ambitious grand-unified and super-unified theories that attempt to unite all forces and particles. One of the most interesting theoretical spin-offs from research into symmetry-breaking has been the realization first that symmetry-breaking invariably produces topological defects and second that, according to the standard Big Bang model, these would have been copiously produced in the early Universe.

The authors of this volume have both made seminal contributions to this field. Alexander Vilenkin was one of the first to suggest that line-like defects ('cosmic strings') could provide a mechanism for the formation of structure in the Universe.

At a phase transition with the appropriate symmetry-breaking, a dense

spaghetti-like tangle of cosmic string would form. As time proceeded, the string would be stretched out by the expansion of the Universe, creating disturbances in the matter on larger and larger scales. Eventually these disturbances would seed the formation of galaxies and clusters of galaxies. As far-fetched as it sounds, this idea is in principle remarkably predictive, providing in the simplest case a complete theory of the formation of structure in the Universe with only one free parameter, the line tension of the string. But detailed predictions have been slow to emerge because the evolution of string networks is a difficult nonlinear problem, requiring computer simulations at the limit of what is currently possible. Paul Shellard is an expert in this area, as well as having made important contributions to the theory of vortons and axion strings. Both authors are distinguished in the scientific literature for the clarity and simplicity of their writing, so it is no surprise to find this book a model of scientific exposition.

The authors begin with an account of phase transitions in theories of high-energy physics, with a nice summary of the topological classification of defects in field theory. The bulk of the book is then devoted to cosmic strings, in all their varieties — superconducting strings, springs, vortons and axion strings. The rest deals with domain walls, monopoles and textures, and the way in which defects might be produced during inflation.

The volume should prove to be an invaluable reference for students and researchers in high-energy theory, providing a nice overview of a fairly large body of work.

The speculative nature of all these ideas can hardly be denied. What has made the field much more interesting however has been the discovery of primordial fluctuations in the cosmic microwave background radiation by the COBE (Cosmic Background Explorer) satellite. Each of the hypothetical defects produced at phase transitions in the early Universe would leave a distinctive signature on the microwave sky — lines for strings, spots for textures — at a level that should allow their measurement by experiments conducted over the next decade. The discovery of any of these defects would be a new landmark for cosmology and particle physics, because it would give us information about physics beyond the standard model, as well as answering one of the most fundamental questions, that of how structure formation began in the Universe. □

Neil Turok is in the High Energy Physics Group, Princeton University, Jadwin Hall, PO Box 708, Princeton, New Jersey 08544, USA.

New Journals

This year, *Nature's* annual new journals review supplement will appear in the issue of 21 September. Publishers and learned societies are invited to submit journals for review, taking note of the following criteria:

- Journals that first appeared during or after June 1993 and issued at least four separate numbers by the end of April 1995 will be considered.
- Journals covering any aspect of science are eligible, although those dealing with clinical medicine and pure mathematics are excluded, as are publications of abstracts.
- Frequency of publication must be at least three times a year. The main language used must be English. Translation journals in English are, of course, eligible.
- Deadline for submission is the end of May.

When submitting journals for review, please send at least four different issues (the first, the most recent and any two others) of each title, together with full details of subscription rates (personal and institutional) and frequency of publication to: Peter Tallack, *Nature*, Macmillan Magazines Ltd, Porters South, Crinan Street, London N1 9SQ, UK. Tel: +44 (0)171 843 4567.

Implications of the late Palaeozoic oxygen pulse for physiology and evolution

Jeffrey B. Graham, Robert Dudley, Nancy M. Aguilar & Carl Gans

The late Palaeozoic was marked by significant changes in atmospheric chemistry and biotic composition. Geochemical models suggest a marked increase and then decline of atmospheric oxygen and associated shifts in the concentration of carbon dioxide. Although the actual magnitude of these changes is uncertain, the pulse of oxygen concentration may have reached a maximum of 35% and then dropped to 15% (compared with the present 21%). This oxygen pulse may have influenced the evolution of major groups of organisms.

MOST models of atmospheric evolution indicate that oxygen fluctuated substantially at various times during the Phanerozoic¹⁻⁶. The recently developed oxygen model of Berner and Canfield⁶ confirms general features of several earlier reconstructions^{2,4}, and indicates a rise in oxygen to 35% and then a fall to 15% over the span of 120 Myr during the late Palaeozoic (Fig. 1a). Although these atmospheric oxygen fluctuations are thought to have been largely driven by biotic factors^{1,7-9}, there has been little consideration of how oxygen changes affected organismic physiology or biosphere evolution.

This essay focuses on the biological effects of hyperoxia and describes changes in the fossil record correlating with varying atmospheric oxygen levels. Relative to the present atmospheric level of 21% oxygen, the Palaeozoic oxygen changes shown in Fig. 1a are substantial and would have had dramatic biological consequences. An increased global oxygen supply during the Mid-Devonian to late Carboniferous would have enhanced diffusion-dependent processes such as respiration. Accordingly, certain organisms could attain a larger body size. An increased aquatic oxygen concentration and greater oxygen penetration into benthic substrates would permit greater biomass densities in most habitats. Elevated oxygen could also increase metabolic rate and turnover and resource accessibility, thereby promoting the radiation of some taxa. Associated atmospheric changes in air density and heat capacity would have affected several biological processes.

In contrast, a decline in global oxygen during the Permian (Fig. 1a) would require compensatory respiratory adaptations and might have been restrictive for some groups. The rate of oxygen decline was too gradual, however, to have been the primary cause of the end-Permian extinction. A 15% oxygen atmosphere at sea level, for example, is equivalent to the oxygen partial pressure (15 kPa in dry air) at an altitude of about 2.5 km, which is well within the adaptation limits for most extant biota¹⁰.

Effects of changing O₂ concentration

Table 1 compares biologically important physical differences between the existing atmosphere and the estimated extremes of the late Palaeozoic oxygen pulse. In plants and most animals, respiration is entirely diffusion limited, thus tissue composition and metabolic demand as well as body shape and thickness are determined by the availability of oxygen¹¹⁻¹³. Table 1 shows oxygen effects on diffusion-dependent body thickness. Relative to dimensions dictated by the present atmospheric level, organisms in 35% oxygen could have a 27% greater diffusive distance, allowing for thicker constructions, less work invested in gas exchange, and greater body size. In contrast, the estimated end-Permian atmosphere containing 15% oxygen would necessitate an 18% reduction in diffusive distance. Biological effects from changes in ultraviolet radiation during the oxygen pulse are unlikely given that all of these atmospheric oxygen levels far

exceed the minimal amounts required for a biologically effective ozone shield^{4,14,15}.

Nitrogen partial pressures are thought to have been fairly constant throughout the Phanerozoic¹⁵⁻¹⁷. Thus, relative to present atmospheric levels, a 35% oxygen atmosphere would have a 21% greater air density and barometric pressure, whereas a 15% oxygen atmosphere would be 13% less (Table 1). A denser atmosphere would favourably influence the evolution of insect flight by offering greater lift and altering characteristics such as the Reynolds number and boundary layer thickness^{18,19}. Density changes would also affect biological processes ranging from ventilatory mechanics to wind-shear resistance. Physical parameters such as viscosity, specific heat and thermal conductivity would also vary in the low and high oxygen atmospheres (Table 1). Calculations of such variation require physical constants based on standard temperature and pressure, however, which may not apply to the variable oxygen conditions of the late Palaeozoic atmosphere.

Palaeoatmospheric models for carbon dioxide^{7,8} show that this gas underwent large and sometimes inverse fluctuations compared to those for oxygen (Fig. 1b). The high carbon dioxide levels occurring in the Ordovician and Silurian resulted from tectonic activity^{4,8,15}. During the late Carboniferous oxygen peak, atmospheric carbon dioxide concentration was similar to PAL (0.036%) but was about three times greater at the end of the Permian. Carbon dioxide shifts of this magnitude affect a number of atmospheric properties (Table 1), aquatic pH, as well as animal and plant physiology (see below).

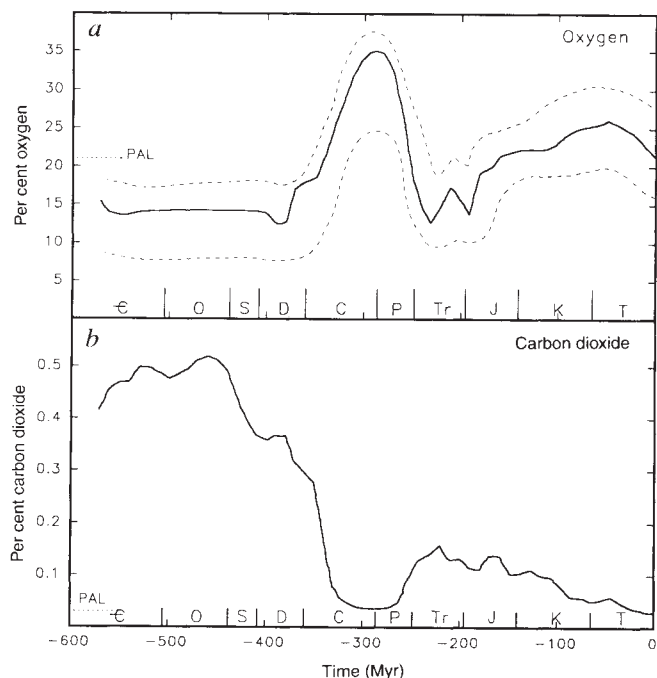
Late Palaeozoic biological correlations

The mid- to late Palaeozoic (Devonian-Permian) saw major shifts in the diversity and complexity of nearly all groups of plants and animals throughout the biosphere²⁰⁻²⁹. The Permian saw a dramatic restructuring punctuated by the end-Permian extinction²³⁻²⁵.

These biotic changes and shifts in divergence of major groups proceeded in parallel with the rise and then fall of oxygen availability. However, the late Palaeozoic was also characterized by changes in physical factors such as climate, tectonics and geochemical cycling^{7,23-25}. Some workers have linked localized oxygen reductions during the Permian to the decline and disappearance of taxa^{9,30}; however, the most likely effect of global atmospheric hypoxia (Fig. 1a) was elimination of species specialized for the permissive hyperoxic atmosphere.

Terrestrial invertebrates. Insects and other arthropod groups provide spectacular examples of the expansive and contractive influences associated with the oxygen pulse. There was a marked Carboniferous diversification of insects on the basis of feeding specializations^{23,25} and the origin of flight, and many species exhibited gigantism relative to extant forms. Predatory defence is the usual explanation offered for insect gigantism³¹. However,

FIG. 1 a, Palaeoatmospheric oxygen concentration model⁶ showing the estimate (solid line) and its range (dashed lines), based upon sensitivity analysis. Between the mid-to-late Devonian (380–360 Myr), oxygen increased from ~18 to 20%, and then rose sharply to ~35% by late Carboniferous (286 Myr). The present atmospheric level (PAL) of 21% is indicated. Oxygen steadily declined throughout the Permian (286–250 Myr) and dropped to about 15% by the end of the Palaeozoic (250 Myr). Model is based on the exchange rate of fixed carbon between the atmosphere, ocean and sediments^{4,6,7} and on various assumptions^{6,9,83}. b, Palaeozoic atmospheric carbon dioxide model^{7–9}. This gas was present in relatively large amounts in the Ordovician–Silurian, fell precipitously during the Devonian–Carboniferous, and increased in the late Permian. The minimum value shown for the late Carboniferous and early Permian approximates PAL for carbon dioxide (about 0.036%; dotted line). Model is based on estimates of fixed carbon exchange and transfer processes within the inorganic, carbonate–silicate cycle^{1,7,8}.



by increasing diffusive permeation, increased oxygen concentration would have permitted insects to become larger. Insects have a tracheal gas-exchange system, a branching network of tubes extending from the body surface to the small tracheoles within respiring cells. This system is primarily, if not exclusively, diffusion dependent, with 0.5 cm suggested as an upper limit to passive diffusion^{32,33}. Gigantism occurred in the Protodonata (dragonflies; wing span up to 71 cm), the Ephemeroptera (mayflies, 9–20 cm),

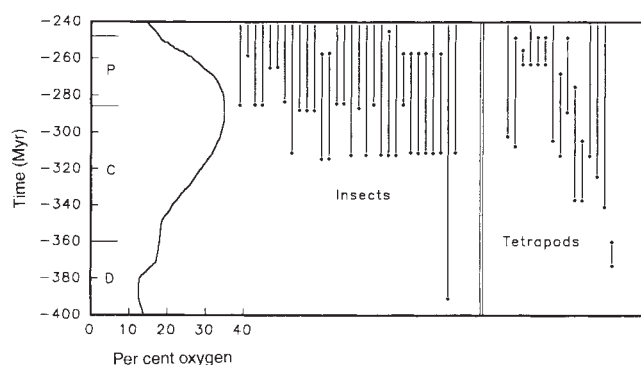
the Palaeodictyoptera (1–43 cm) and in the Diplura and Thysanura^{34–37}. For example, the giant Carboniferous dragonfly *Meganeura monyi* had a thoracic diameter of about 2.8 cm³⁸. By contrast, thoracic width of the largest (10-cm wingspan) extant dragonflies (*Anax*, *Aeshna*) is about 1 cm^{38,39} (M. May, personal communication). Most of the various insect taxa that attained exceptionally large body sizes during the Carboniferous did not persist after the Permian, when 27–30% of the known insect orders were lost (Fig. 2)^{40–42}. Gigantism also characterized some

TABLE 1 Comparison of physical properties of the present oxygen atmosphere (21% O₂) with those of the relatively hyperoxic late Carboniferous (35% O₂) and the relatively hypoxic end-Permian (15% O₂)

	21% O ₂ (present)	35% O ₂ (285 Myr)	15% O ₂ (250 Myr)	Biological importance
Respiratory gases				
Oxygen				
O ₂ partial pressure (kPa)	21.2	35.3	15.1	Respiration, lignin biosynthesis
Krogh's maximum radius (cm)*	0.11	0.14	0.09	Size limit for diffusion dependence
Water O ₂ content (ml l ⁻¹)†	6.9	7.4	4.9	Aquatic respiration
Carbon dioxide‡				
CO ₂ partial pressure (kPa)	0.03	0.03	0.09	Effects on photosynthesis, moisture content and global energy balance
Water CO ₂ content (ml l ⁻¹)	0.31	0.31	0.9	Aquatic pH effects, acid–base balance and ion regulation
Air properties				
Density (kg m ⁻³)	1.29	1.56§	1.12§	Flight and respiratory mechanics, wind shear
Dynamic viscosity kg (m ⁻¹ s ⁻¹)	18.2 × 10 ⁻⁶	+	–	Boundary layer thickness
Specific heat (Jg ⁻¹ deg ⁻¹)	1.006	+	–	Heat capacity and relative humidity
Thermal conductivity (Js ⁻¹ m ⁻¹ deg ⁻¹)	2.4 × 10 ⁻²	+	–	Earth thermal budget, climate

For values not calculated, the plus and minus symbols represent an increase and decrease, respectively, relative to the present 21% oxygen atmosphere.
* The equation for the maximum radius of a diffusion-dependent animal is $r \leq (6KpO_2/vO_2)^{-2}$, where pO_2 is oxygen partial pressure at its surface, vO_2 is oxygen-consumption rate, K is Krogh's diffusion constant, and 6 is a geometric shape constant^{13,14}.
† Values calculated for fresh water at 20 °C and constant pH.
‡ Palaeozoic CO₂ changes partly in and out of phase with those for O₂ (see Fig. 1 and text).
§ Calculation assumes nitrogen and trace gas partial pressures similar to present atmosphere.

FIG. 2 Stratigraphic ranges of 30 insect orders⁴¹ and 16 tetrapod morphotype lineages⁶⁸ in relation to the late Palaeozoic oxygen pulse⁶. Record extends from the Lower Devonian (400 Myr) to the Lower Triassic (240 Myr). The Carboniferous diversification of both the insects and vertebrates correlates with the rise in atmospheric oxygen. A correlation also exists for the end-Permian (250 Myr) disappearance of members of both groups and the fall in atmospheric oxygen. Insect data are from ref. 41 modified according to ref. 37 (see also refs 40 and 84). Estimates for insect extinctions at the close of the Permian range from 8 of 30 orders (27%)⁴¹ to 8 of 27 orders (30%)⁴⁰. The 75% extinction rate shown for end-Permian tetrapods in this figure parallels other estimates (67% of amphibian families and 78% of reptilian families²⁵).



other Carboniferous arthropods (such as diplopods, arthropleuroids and scorpions)^{26,36,43-46}, extant relatives of which are diffusion dependent⁴⁷.

The greater density of the hyperoxic atmosphere facilitated insect flight by enabling relatively small articulated winglets to provide lift⁴⁶, albeit with suboptimal aerodynamics. Although the Devonian may have seen the first apterygote insects⁴⁸, winged forms diversified extensively throughout the Carboniferous^{40,49}. The earliest proto-wings possibly had a respiratory function and were used secondarily for locomotion and predator escape^{39,41}. The need to power flight would have required ever-more sustained and higher levels of oxidative metabolism, which would be facilitated by higher levels of atmospheric oxygen. The ability of some Recent flying insects to ventilate the tracheal system minimizes diffusive limitations; however, mechanical ventilation does not mitigate diffusion limitation in the terminal tracheoles^{39,47}. Schidlowski⁵⁰ argued that the tracheal system of *Meganeura* would have been inadequate for sustaining flight metabolism in oxygen at present atmospheric levels, and predicted a hyperoxic Carboniferous atmosphere.

Flight became the decisive agent for insect dispersal and diversification. The improvement of flight and its association with smaller species may have ensured its persistence in numerous pterygote taxa through the Permian. But as already noted, large fliers did not survive beyond the Permian^{37,40}.

Aquatic invertebrates. The majority of aquatic invertebrates lack well-developed respiratory and ventilatory systems and depend upon diffusion-mediated respiration^{12,51}. Increased oxygenation of Devonian and Carboniferous fresh waters would have increased their utility as nursery grounds for eggs and larvae, thereby expanding life-history options for the rapidly diversifying terrestrial arthropods. Beyond this, global atmospheric hyperoxia would have reduced the limitation of seasonal aquatic hypoxia, favoured infaunal penetration of marine sediments, allowed denser aggregations, and increased the body size of some groups.

Size increases have been recorded for a number of taxa throughout geological time, and have been proposed to reflect multiple biotic factors^{27,52,53}. The observed correlation between size increases in marine invertebrates (bryozoans⁵⁴, foraminiferans²⁷, rugose corals²⁷ and brachiopods^{26,27}) and atmospheric hyperoxia is notable because these forms have a primarily diffusion-dependent respiration. Specific examples include the largest known brachiopod (*Gigantoproductus*, with a shell 300 mm wide) of the early Carboniferous, and the radiation of the large fusulinid Foraminifera (Mid-Carboniferous–Permian).

The evolutionary transition from a 'brachiopod-dominated late Palaeozoic marine fauna' to the modern, 'mollusc-dominated Mesozoic–Cenozoic fauna' occurred at the end of the Permian⁵⁵. Associated with this was the end-Permian extinction of several dominant members of the sessile, epifaunal filter-feeding community of the Palaeozoic, including the rugose and tabu-

late corals, and numerous families of brachiopods, bryozoans, crinoids and others. A number of factors, including changes in sea level and reduced temperature, may have contributed to this extinction and faunal transition^{24,25}. Whereas we do not regard Permian reductions in global oxygen as a major causal factor in either of these events, hypoxia probably contributed to the decline of many of the diffusion-dependent faunal elements. The bivalves, which had a relatively greater Permian survival than did the brachiopods, and which became dominant in the Triassic⁵⁶, had numerous specializations that allowed ventilation, mobility and sediment penetration^{26,57,58}.

The terrestrial flora. The Carboniferous and early Permian flora included ferns and primitive gymnosperms, as well as sphenop-sids and lycopods⁵⁹⁻⁶⁴, which were arborescent and grew to considerable height²³. Increased plant height and arborescence required thicker support structures, particularly as a greater atmospheric density elevated wind stress. Heightened atmospheric oxygen facilitated these morphological changes by allowing diffusion through thicker support elements and by enhancing the oxygen-dependent⁵⁹ biosynthesis of lignin, a dominant structural material of many Carboniferous plants^{7,23,64}. An increased oxygen level may not, however, have been entirely advantageous for plants, particularly when it occurred simultaneously with a decline in carbon dioxide (Fig. 1b). Changes in the carbon dioxide to oxygen ratio can lessen plant productivity by photorespiration⁶⁵. However, the rarefaction of atmospheric carbon dioxide over the Carboniferous may have established a selective advantage for new, more effective carbon-fixation pathways in some plants⁶⁶.

Terrestrial vertebrates. The diversification of tetrapods began in the Devonian and was accelerated during the Carboniferous, when at least eleven of sixteen known basal lineages (including the three primary classes of amniotes: Synapsida, Diapsida and Anapsida) all appeared^{23, 25,67,68} (Fig. 2). In the Permian, 75 per cent of these tetrapod lineages went extinct^{25,67,68} (Fig. 2).

Global atmospheric hyperoxia possibly aided the vertebrate invasion of land². Breathing hyperoxic air reduces the ratio of evaporative water loss to oxygen uptake, and thus increases aerial gas exchange efficiency by lowering desiccation⁴⁷. Although the early tetrapods had both gills and lungs⁶⁹⁻⁷¹, gills are not effective for aerial gas exchange, particularly carbon dioxide discharge, which is a more acute constraint for aerial respiration than oxygen uptake^{47,71}. As many early tetrapods were sizeable and had a thick body armour⁷², aerial cutaneous respiration could not compensate for the reduced respiratory capacity of gills in air. Tetrapods were thus heavily dependent on lung exchange^{22,67,72-74}. The rise in atmospheric oxygen coupled with the reduction in carbon dioxide would have elevated primitive lung effectiveness. Subsequent changes in these gas ratios could also have influenced the transition from pulse pumping to aspiration breathing and perhaps even the shift from an oxygen to a carbon dioxide-modulated respiratory control mechanism. The reduced ratio of water loss to oxygen uptake achievable in

hyperoxic air may have also been critically important in the evolution of the cleidoic egg.

Hyperoxia would have also enhanced tetrapod metabolic capacity and thus contributed to the sustained power production required to overcome gravity. Increased metabolic performance opened new ecological options within the rapidly expanding terrestrial biosphere. Although tetrapod diversity decreased in the Permian, the synapsids underwent a pronounced diversification, proceeding from the pelycosaurs to the therapsids, a diverse assemblage of herbivores and carnivores^{22,24,67,74} known as the mammal-like reptiles²². The diversification of synapsids seems to be partially attributable to the effects of hyperoxia and a denser atmosphere on activity enhancing specializations such as metabolic heat retention. The large 'sails' of pelycosaurs such as *Dimetrodon* are thought to have functioned in heat transfer^{22,74}. The capacity to regulate heat gain and loss was an important precursor for endothermy⁷⁵.

Conclusion

Attempts to correlate atmospheric oxygen with biosphere evolution have included the origin of the ultraviolet-ozone shield and the early radiation of plants^{1,4,14,15}, the evolution of the diffusion-limited Ediacaran fauna⁷⁶⁻⁷⁹, the Cambrian explosion⁸⁰, the correlation of patterns of extinction and metabolic requirements

among various taxa^{2,81}, and regional or even mass extinction events^{9,30,82}.

We propose that a pulse in atmospheric oxygen spanning the mid-Devonian, Carboniferous and Permian periods influenced the contemporaneous biosphere. The rise in oxygen during the Devonian and Carboniferous favourably influenced diffusion-dependent features. Increased oxygen availability may have also fuelled the diversification and ecological radiation of late Palaeozoic groups by acting as a substrate for the evolution of behavioural, physiological and ecological adaptations, permitting greater exploitation of aquatic habitats and the newly evolving terrestrial biosphere. The fall in global atmospheric oxygen during the Permian was restrictive and modified or eliminated some taxa that radiated in hyperoxia, but was not a primary cause of the end-Permian extinction. □

Jeffrey B. Graham and Nancy M. Aguilar are at the Marine Biology Research Division & Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093-0204, USA; Robert Dudley is in the Department of Zoology, University of Texas, Austin, Texas 78712, USA and the Smithsonian Tropical Research Institute, PO Box 2072, Balboa, Republic of Panama; Carl Gans is in the Department of Biology, University of Michigan, Ann Arbor, Michigan 48109, USA.

- Kasting, J. F. *Science* **259**, 920-926 (1993).
- Budyko, M. I., Ronov, A. B. & Yanshin, A. L. *History of the Earth's Atmosphere* (Springer, Berlin, 1985).
- Tappan, H. in *Molecular Oxygen in Biology* (ed. Hayaishi, O.) 81-135 (North Holland, Amsterdam, 1974).
- Berkner, L. V. & Marshall, L. C. *J. Atmos. Sci.* **22**, 225-261 (1965).
- Hudson, J. D. *J. geol. Soc. Lond.* **146**, 155-160 (1989).
- Berner, R. A. & Canfield, D. E. *Am. J. Sci.* **289**, 333-361 (1989).
- Robinson, J. M. *Global Planetary Change* **97**, 51-62 (1991).
- Berner, R. A. *Science* **261**, 68-70 (1993).
- Berner, R. A. *Nature* **340**, 603-604 (1989).
- Bouvierot, P. *Adaptation to Altitude Hypoxia in Vertebrates* (Springer, Berlin, 1985).
- Krogh, A. *The Comparative Physiology of Respiratory Mechanisms* (Univ. Pennsylvania Press, Philadelphia, 1941).
- Graham, J. B. *Am. Zool.* **28**, 1031-1045 (1988).
- Graham, J. B. *Am. Zool.* **30**, 137-146 (1990).
- Kasting, J. F. *Precambrian Res.* **34**, 205-229 (1987).
- Holland, H. D. *The Chemical Evolution of the Atmosphere and Oceans* (Princeton Univ. Press, Princeton, 1984).
- Hart, M. H. *Icarus* **33**, 23-39 (1978).
- Wageneck, P. *Chemistry of the Natural Atmosphere* (Academic, San Diego, 1988).
- Volpe, S. *Life in Moving Fluids* (Princeton Univ. Press, Princeton, 1994).
- Pennycook, C. J. *Newton Rules Biology* (Oxford Univ. Press, Oxford, 1992).
- Behrensmeyer, A. K. (ed.) *Terrestrial Ecosystems through Time: Evolutionary Paleocology of Terrestrial Plants and Animals* (Chicago Univ. Press, Chicago, 1992).
- Niklas, K. L., Tiffney, B. H. & Knoll, A. H. in *Phanerozoic Diversity Patterns: Profiles in Macroevolution* (ed. Valentine, J. W.) 97-128 (Princeton Univ. Press, Princeton, 1985).
- Carroll, R. L. *Vertebrate Paleontology and Evolution* (Freeman, New York, 1988).
- Shear, W. A. *Nature* **351**, 283-289 (1991).
- Erwin, D. H. *The Great Paleozoic Crisis: Life and Death in the Permian* (Columbia Univ. Press, New York, 1993).
- Erwin, D. H. *Nature* **367**, 231-236 (1994).
- Clarkson, E. N. K. *Invertebrate Paleontology and Evolution* (Chapman & Hall, London, 1993).
- Newell, N. D. *Evolution* **3**, 103-124 (1949).
- Little, C. *The Colonisation of Land: Origins and Adaptations of Terrestrial Animals* (Cambridge Univ. Press, Cambridge, 1990).
- Panchen, A. L. (ed.) *The Terrestrial Environment and the Origin of the Land Vertebrates* (Academic, London, 1980).
- Holser, W. T., Schönlaub, H. P., Boeckelmann, K. & Magaritz, M. *Abh. Geol. Bundesanst.* **45**, 1-232 (1991).
- Vermeij, G. J. *Evolution and Escalation* (Princeton Univ. Press, Princeton, 1987).
- Weis-Fogh, T. *J. exp. Biol.* **41**, 229-256 (1964).
- Mill, P. J. in *Comprehensive Insect Physiology, Biochemistry and Pharmacology* Vol. 3 (eds Kerkut, G. A. & Gilbert, L. I.) 518-593 (Pergamon, Oxford, 1985).
- Kukalová-Peck, J. *Can. J. Zool.* **63**, 933-955 (1985).
- Kukalová-Peck, J. *Can. J. Zool.* **65**, 2327-2345 (1987).
- Shear, W. A. & Kukalová-Peck, J. *Can. J. Zool.* **68**, 1807-1834 (1990).
- Carpenter, F. M. *Treatise on Invertebrate Paleontology* Vols 3, 4 (Univ. Kansas Press, Lawrence, 1992).
- May, M. L. *Evolution* **36**, 1051-1058 (1982).
- Heinrich, B. *The Hot-blooded Insects* (Harvard Univ. Press, Cambridge, 1993).
- Labandeira, C. C. & Sepkoski, J. J. *Jr Science* **261**, 310-316 (1993).
- Kukalová-Peck, J. in *Insects of Australia* (eds Naumann, I. D. et al.) 141-179 (Cornell Univ. Press, Ithaca, 1991).
- Benton, M. J. (ed.) *The Fossil Record 2* (Chapman and Hall, London, 1993).
- Kraus, O. *Zool. Soc. Lond. Symp.* **32**, 13-22 (1974).
- Rölfe, W. D. I. in *The Terrestrial Environment and the Origin of Land Vertebrates* (ed. Panchen, A. L.) 117-157 (Academic, London, 1980).
- Briggs, D. E. G. in *Special Papers in Paleontology* **33**, 157 (1985).
- Ellington, C. P. *Adv. Ins. Physiol.* **23**, 171-210 (1991).
- Withers, P. C. *Comparative Animal Physiology* (Saunders, New York, 1992).
- Labandeira, C. C., Beall, B. S. & Hueber, F. M. *Science* **242**, 913-916 (1988).
- Wootton, R. J. in *Major Evolutionary Radiations* (eds Taylor, P. D. & Larwood, G. P.) 187-208 (Clarendon, Oxford, 1990).
- Schidlowski, M. *Geol. Rundsch.* **60**, 1352-1384 (1971).
- Mill, P. J. *Respiration in the Invertebrates* (Macmillan, London, 1972).
- Gould, S. J. *Biol. Rev.* **41**, 587-640 (1966).
- McKinney, M. L. in *Evolutionary Trends* (ed. McNamara, K. J.) 75-118 (Univ. Arizona Press, Tucson, 1990).
- Elias, M. K. *Paleontology* **2**, 306-336 (1937).
- Sepkoski, J. J. *Jr Paleobiology* **7**, 36-53 (1981).
- Sepkoski, J. J. Jr & Hulver, M. L. in *Phanerozoic Diversity Patterns: Profiles in Macroevolution* (ed. Valentine, J. W.) 11-39 (Princeton Univ. Press, Princeton, 1985).
- Sepkoski, J. J. *Jr Paleobiology* **10**, 246-267 (1984).
- Erwin, D. H. *Trends Ecol. Evol.* **4**, 225-229 (1989).
- Chapman, D. J. in *Geological Factors and the Evolution of Plants* (ed. Tiffney, B. H.) 23-45 (Vail-Ballou, New York, 1985).
- Stewart, W. N. *Paleobotany and the Evolution of Plants* (Cambridge Univ. Press, 1983).
- Scott, A. C. in *The Terrestrial Environment and the Origin of Land Vertebrates* (ed. Panchen, A. L.) 87-115 (Academic, London, 1980).
- Tiffney, B. H. in *Geological Factors and the Evolution of Plants* (ed. Tiffney, B. H.) 1-10 (Vail-Ballou, New York, 1985).
- Tiffney, B. H. in *Paleobotany, Paleocology and Evolution* (ed. Niklas, K. J.) 193-230 (Praeger, New York, 1981).
- Robinson, J. M. *Geology* **15**, 607-610 (1990).
- Lawlor, D. W. *Photosynthesis: Metabolism, Control and Physiology* (Wiley, New York, 1987).
- Ehleringer, J. R. & Monson, R. K. *Rev. Ecol. Syst.* **24**, 411-439 (1993).
- Carroll, R. L. *Rev. Earth. planet. Sci.* **20**, 45-84 (1992).
- Lombard, R. E. & Sumida, S. S. *Am. Zool.* **32**, 609-622 (1992).
- Thomson, K. S. *Am. Sci.* **79**, 488-490 (1991).
- Coates, M. I. & Clack, J. A. *Nature* **352**, 234-236 (1991).
- Gans, C. *Forma et Functio* **3**, 61-104 (1970).
- Romer, A. S. *Resp. Phys.* **14**, 183-192 (1972).
- Bray, A. A. *Phil. Trans. R. Soc. Lond.* **309**, 289-322 (1985).
- Colbert, E. H. & Morales, M. *Evolution of the Vertebrates* (Wiley-Liss, New York, 1991).
- Hammel, H. T. *J. therm. Biol.* **8**, 37-42 (1983).
- Cloud, P. *Paleobiology* **2**, 351-387 (1976).
- Runnegar, B. *Alcheringa* **6**, 223-239 (1982).
- Runnegar, B. *J. geol. Soc. Aust.* **29**, 395-411 (1982).
- Knoll, A. H. et al. *Nature* **231**, 832-838 (1986).
- McMenamin, M. A. S. & McMenamin, D. L. S. *The Emergence of Animals-The Cambrian Breakthrough* (Columbia Univ. Press, New York, 1990).
- McAlester, A. L. *J. Paleont.* **44**, 405-409 (1971).
- Wignall, P. B. & Hallam, A. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **93**, 21-46 (1992).
- Jones, T. P. & Chaloner, W. G. *Global Planetary Change* **97**, 39-50 (1991).
- Ross, A. J. & Jarzembowski, E. A. in *The Fossil Record 2* (ed. Benton, M. J.) 363-426 (Chapman and Hall, London, 1993).

ACKNOWLEDGEMENTS. We thank R. Burnham, R. Carr, R. L. Carroll, P. A. Fields, D. C. Fisher, P. Gingerich, M. S. Gordon, R. R. Hessler, K. C. Lohmann, F. Posell, M. L. May, W. A. Newman, R. H. Rosenblatt, H.-P. Schultze and S. Sumida, for advice, information and discussion about the diverse subjects brought together in this paper, and for reading and commenting on manuscript drafts. This study was supported in part by both NSF and the Vetlesen Foundation (J.B.G.), a Reeder Fellowship (R.D.), a Ford Foundation Fellowship (N.M.A.), and the Leo Leeson Foundation for Tropical Biology (C.G.).

Are we aware of neural activity in primary visual cortex?

Francis Crick* & Christof Koch†

* The Salk Institute, 10010 North Torrey Pines Road, La Jolla, California 92037, USA

† Computation and Neural Systems Program, 139-74, California Institute of Technology, Pasadena, California 91125, USA

It is usually assumed that people are visually aware of at least some of the neuronal activity in the primary visual area, V1, of the neocortex. But the neuroanatomy of the macaque monkey suggests that, although primates may be aware of neural activity in other visual cortical areas, they are not directly aware of that in area V1. There is some psychophysical evidence in humans that supports this hypothesis.

WHAT is the relationship between 'awareness', in particular visual awareness, and neuronal activity in the nervous system? Within a larger context, this question is sometimes known as the 'mind-body' problem, and has been asked since antiquity. In recent years, we have been trying to understand this relationship from a neuronal, reductionist point of view by making the following two basic assumptions.

(1) For an animal to be aware of some aspect of a visual object, there has to be a group of similar neurons (to allow coarse coding) located somewhere in the cortical system (which includes the neocortex and the hippocampus, as well as directly associated parts, such as the thalamus), the firing of which is correlated with that feature in the visual scene. We shall not discuss here the so-called binding problem (that is, how different aspects of a single visual object are bound together), nor the idea that this binding may require correlated firings (such as the γ oscillations) of sets of neurons in different cortical areas or the detailed nature of the neuronal correlate of consciousness (NCC)¹⁻⁵.

(2) Such neurons must project directly to some part of the front of the cortex, in particular, to those areas in front of the primary motor area (M1, also called area 4). In general, this large expanse of cortex is called 'frontal' cortex and includes motor cortex, premotor and prefrontal areas.

We have argued elsewhere⁵⁻⁷ that to be aware of an object or event the brain has to construct a multilevel, explicit, symbolic interpretation of part of the visual scene. By multilevel we mean, in psychological terms, different levels such as those that correspond, for example, to lines or to eyes or to faces. In neurological terms we mean, loosely, the different levels in the visual hierarchy⁸.

What exactly do we mean by 'explicit'? This is not an easy question to answer. The pattern of coloured dots on a television screen, for instance, may contain an implicit representation of a person's face, but only the dots and their location are explicitly represented on the screen. A simple neural network (such as a one-layer perceptron) could be constructed to respond whenever a red dot appeared at one particular place on the screen, or even to a particular fixed set of dots. Such a simple network cannot be tuned up to respond to a particular face that might appear anywhere on the TV screen, and varied in size, orientation or facial expression.

It is tempting to say that an explicit representation of a feature means that in the brain there is only a single neuron (otherwise called a 'grandmother' cell) that responds to that particular feature, and to that feature alone. We think this very improbable, if only because the firing of a single neuron in the cortical system is, by itself, unlikely to activate strongly any of its postsynaptic targets. Furthermore, 'normal' visual input is unlikely to activate such a cell by itself without also activating its neighbours.

We believe that any particular visual feature will be coarse coded (that is, spread out over a set of similar neurons), and indeed there is some evidence for this in the case of faces⁹. These 'face neurons' do not respond to a single spot of light (as retinal ganglion cells do) but are activated mainly by face-like objects, each neuron of the set being activated in a somewhat different way. This type of firing of a set of neurons is considered to be the neural symbol for a face, and the collective firing of this set makes the face-like aspect of the face explicit. (This use of the term 'symbol' should not be taken to imply the existence of a disembodied homunculus who is looking at it.) The meaning of such a symbol depends not only on the receptive fields of these neurons but also to which other neurons they project (their projective fields), and exactly how they are connected to them.

The firing of a single retinal ganglion cell might represent explicitly a particular spot of light in the visual field, but could not represent a face. An explicit face representation can only occur at a higher level in the visual system, after much further processing, especially as the 'face-neurons' are found experimentally to be relatively indifferent to the location and orientation of the face in the visual field¹⁰; if they were not there would have to be too many of them. Note that a person is not necessarily aware of all such representations. Something more is required for awareness: see (2) above.

We have not previously discussed in detail the need for projections from the visual system to the front of the brain, although we have mentioned it in passing⁷. Our basic argument assumes that, in going from one visual area to another further up in the visual hierarchy (that is, further away from the retina^{8,11,12}), the information is recoded at each step. This is broadly compatible with the known fact that the 'features' to which a neuron responds become more complex in going from the primary visual cortex, V1 (also called striate cortex, or area 17) to the higher levels in the visual hierarchy, such as the inferotemporal cortical areas^{13,14}.

A neuron that is firing, but whose axon has been inactivated by local anaesthetics, can contribute little, if anything, to immediate visual awareness, unless one favours some dualistic position¹⁵. Thus the main destination of the axon—loosely, its projective field—is clearly important.

Prefrontal brain areas and planning

Our second assumption is based on the broad idea of the biological usefulness of visual awareness (or, strictly, of its neural correlate). This is to produce the best current interpretation of the visual scene, in the light of past experience either of ourselves or of our ancestors (embodied in our genes), and to make it available, for a sufficient time, to the parts of the brain that contemplate, plan and execute voluntary motor outputs (of one sort or another).

Exactly how these prefrontal and premotor cortical areas operate is currently unknown, although there is now fragmentary evidence about the behaviour of some of them. Even in the macaque, the details of neuroanatomical connections between all these areas have not yet been worked out in as much detail as they have for most of the visual areas¹⁶⁻¹⁹.

It is probably a general rule that the further, in terms of the number of stages, a prefrontal area is from the primary motor area M1, the longer is the timescale of the planning in which it is engaged^{17,20}. Moreover, these cortical areas are all heavily involved with the basal ganglia (which include the neostriatum, the globus pallidus and the substantia nigra), the main function of which may be to provide a bias back to these areas (as well as to the superior colliculus in the midbrain) to influence the next step in their processing, that is, to assist some behaviours that involve a sequence of activities. The subject is additionally complicated for humans because of our highly developed language system and its usefulness for expressing our 'thoughts' (in silent speech, for example).

Fortunately, at this stage, the details of the behaviour of these 'frontal' areas need not concern us. All we need to postulate is that, unless a visual area has a direct projection to at least one of them, the activities in that particular visual area will not enter visual awareness directly, because the activity of frontal areas is needed to allow a person to report consciousness.

Primary visual cortex and its connections

This lack of any direct projection to the frontal cortex appears to be true for area V1 of the macaque monkey, the almost exclusive recipient of the output of the lateral geniculate nucleus (LGN). V1 has no direct projections to the frontal eye fields (part of area 8), nor to the broad prefrontal region surrounding and including the principal sulcus (see Table 3 in ref. 8), nor, as far as we know, to any other 'frontal' area (see Fig. 1). Nor does it project to the caudate nucleus of the basal ganglia²¹, to the intralaminar nuclei of the thalamus (L. G. Ungerleider, personal communication), the claustrum²² or the brainstem, with the exception of a small projection from peripheral V1 to the pons²³. V1 does, of course, provide the dominant visual input to most of the posterior visual cortical areas, including V2, V3, V4 and MT (also known as V5¹¹). Among subcortical targets, V1 projects to the superior colliculus²⁴, the LGN and the pulvinar^{25,26}.

It is unlikely that information sent along the pathway from V1 to the superior colliculus (partly responsible for initiating and controlling eye movements) can produce conscious visual awareness. There is a multistage pathway going from V1 to the colliculus, from there to the pulvinar (a large, mainly visual, part of the thalamus) and thence to higher visual areas. This pathway may be involved in visual attention²⁶, but, according to our arguments, it is not sufficiently direct or strong to produce, by itself, vivid visual awareness of the neural activities in V1.

The pathway from V1 to the colliculus might possibly be used to produce involuntary eye movements, so psychophysical tests, using eye movements as the response, might show a form of blindsight^{27,28} in which subjects respond above chance while denying that they see anything. It is also possible that this or other pathways can produce vague feelings of some sort of awareness.

Primary visual cortex and awareness

Our hypothesis is too speculative to be convincing as it stands, as we are not yet confident as to how to think correctly about most of the operations of the brain, and especially about the detailed function of the so-called 'back pathways'. Some readers may find our suggestion counterintuitive, partly because for many years V1 was the only visual cortical area that was worked on extensively. We would ask them: do you believe that you are directly aware of the activity in your retina? (Of course, without your retinae, you cannot see anything.) If you do not believe

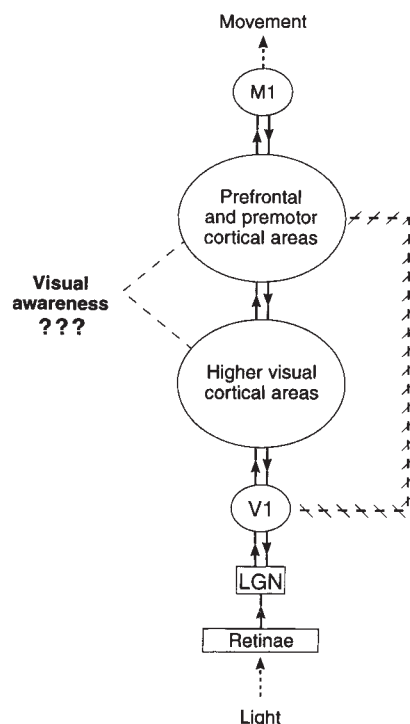


FIG. 1 An extremely schematic diagram of some of the brain areas involved. V1 is the primary visual cortex (also called the striate cortex or cortical area 17). M1 is the primary motor cortex (also called cortical area 4). LGN is the lateral geniculate nucleus of the thalamus. The main point of this paper is that the connection shown hatched on the right of the diagram is missing in the macaque monkey. Many relevant areas, such as the rest of the thalamus, the superior colliculus, the basal ganglia, the cerebellum, the brain stem and the claustrum, have been omitted for simplicity, as have the other sensory systems. The question marks indicate that, at the moment, we are unsure of the exact nature of visual awareness, although we argue here that it does not directly involve neurons in V1.

this, what is the argument that you are directly aware of the neural activity in V1?

To avoid misunderstanding, let us underline what our hypothesis does not say. We are not suggesting that the neural activity in V1 is unimportant. On the contrary, we believe the detailed processing in V1 is highly important for normal vision, although recent work²⁹ has shown that V1 in at least one patient is not essential for some limited form of visual awareness related to motion perception. All we are hypothesizing is that the activity in V1 does not directly enter awareness. What does enter awareness, we believe, is some form of the neural activity in certain higher visual areas, because they project directly to prefrontal areas. This seems well established for cortical areas in the fifth tier of the visual hierarchy, such as MT and V4. For areas in the intervening tiers, such as V2, V3, V3A, VP and PIP, we prefer to leave the matter open for the moment (see Table 3 in ref. 8).

Experiments suggest that only some of the active neurons in a cortical area are likely to produce direct visual awareness. During binocular rivalry^{30,31}, the visual input is constant but the visual percept changes. Experiments on neurons in cortical area MT of the alert macaque monkey show that during binocular rivalry produced by two gratings moving in opposite directions, only a subset of the active neurons in cortical area MT follow the percept³². It will be of great importance to discover the layer in which these cells are located, their type and where they send their axons. The firing of most of the neurons in MT does not

depend on which of the two alternative percepts the monkey reports seeing.

Our hypothesis was suggested by neuroanatomical data from the macaque monkey. For humans, we are less certain because of the present miserable state of human neuroanatomy³³, but we surmise that our hypothesis, if true for the macaque monkey, is also likely to be true for apes and humans. To be established as correct, it also needs to fit with all the neurophysiological and psychological data. What kind of evidence would support it?

Physiological and psychophysical evidence

A possible example may make this clearer. It is well known that the colour perceived at one particular visual location is influenced by the wavelengths of the light entering the eye from surrounding regions in the visual field^{34,35}. This mechanism acts partially to compensate for the effects of differently coloured illumination. A white patch surrounded by patches of many colours still looks fairly white even when illuminated by pink light. This form of (partial) colour constancy is often called the Land effect³⁴.

It has been shown in the anaesthetized monkey³⁶⁻³⁸ that neurons in V4, but not in V1, exhibit the Land effect. As far as we know, the corresponding information is lacking for alert monkeys. Because people cannot voluntarily turn off the Land effect, it would follow, if the same results could be obtained in a behaving monkey, that it would not be directly aware of the 'colour' neurons in V1. Notice that if neurons in both V1 and V4 in the alert monkey did turn out to show the full Land effect, this would not, by itself, disprove our hypothesis, as we do believe that people are visually aware of certain neural activity in V4 that could be triggered by activity in V1.

Psychophysical experiments would support our hypothesis if they demonstrate that people are not aware of neuronal activity that is highly likely to occur in V1. Such experiments have been done recently by D. I. MacLeod and Sheng He (personal communication). In brief, they have shown that exposure to high-contrast gratings that are so finely spaced that they cannot be seen (that is, cannot be distinguished from a uniformly grey surface) can produce an orientation-selective loss in sensitivity of human subjects to slightly less finely spaced gratings that can indeed be perceived. Because of neuronal convergence in higher cortical areas and the associated increase in receptive field size, neurons sensitive to the very fine conditioning grating appear to be restricted to V1 (refs 39, 40). (This is one of the reasons why it is frequently assumed that we are aware of neural activity in V1.) These psychophysical experiments are therefore compatible with the idea that certain neurons in V1 respond to very high

spatial frequencies of which we are not visually aware. (We did not know of these results when we first formulated our hypothesis.) The support for our ideas would be greater if it were shown (by imaging methods such as positron-emission tomography (PET) or functional magnetic resonance imaging (MRI)) that these invisible gratings produced significant activity in V1 in humans. For an alert macaque, it might be possible to show experimentally that very finely spaced gratings, that activated certain neurons in V1, could not be reported by the monkey, although the animal could report less finely spaced ones.

The firing of some neurons in V1 depends upon which eye the visual signal is coming through. Neurons higher in the visual hierarchy do not make this distinction, that is, they are typically binocular. Most people are certainly not vividly and directly aware of which eye they are seeing with (unless they close or obstruct one eye), although whether they have some very weak awareness of the eye of origin is more controversial⁴¹. These well-known facts suggest that people are not vividly aware of much of the activity in V1.

Our ideas would not be disproved if it were shown convincingly that (for some people) V1 is activated during visual imagery tasks (for debate, see ref. 42). There is no obvious reason why such top-down effects should not reach V1. Such V1 activity would not by itself prove that people are directly aware of it, any more than the V1 activity produced there when their eyes are open proves this.

Theories are rarely proved or disproved by a single experiment; it needs a coherent body of interlocking data to establish a theory as correct. It could be argued, in opposition to our view, that there is indeed a small subset of V1 neurons the firing of which directly expresses some unique aspect of visual awareness. These special neurons could easily be missed, making it difficult to disprove our idea. If it turns out that in many other cortical areas only a particular type of neuron expresses visual awareness (for example, some of those in the lower cortical layers), and that neurons of a similar type in V1 respond in ways of which we are not aware, this might be sufficient proof of our hypothesis.

Our hypothesis is thus rather subtle; although we believe that if it turns out to be true it will eventually come to be regarded as completely obvious. We hope that further neuroanatomical work will make it plausible for humans, and further neurophysiological studies will show it to be true for most primates. We have yet to track down the location and nature of the neural correlates of visual awareness⁴³. Our hypothesis, if correct, would narrow the search to areas of the brain further removed from the sensory periphery. □

Received 19 December 1994; accepted 25 February 1995.

1. Milner, P. M. *Psychol. Rev.* **81**, 521-535 (1974).
2. von der Malsburg, C. & Schneider, W. *Biol. Cybern.* **54**, 29-40 (1986).
3. Eckhorn, R. et al. *Biol. Cybern.* **60**, 121-130 (1988).
4. Gray, C. M., König, P., Engel, A. K. & Singer, W. *Nature* **338**, 334-337 (1989).
5. Crick, F. & Koch, C. *Semin. Neurosci.* **2**, 263-275 (1990).
6. Crick, F. & Koch, C. *Scient. Am.* **267**, 152-159 (1992).
7. Koch, C. & Crick, F. in *Large-scale Neuronal Theories of the Brain* (eds Koch, C. & Davis, J.) 93-109 (MIT Press, Cambridge, Massachusetts, 1994).
8. Felleman, D. J. & Van Essen, D. *Cerebr. Cortex* **1**, 1-47 (1991).
9. Young, M. P. & Yamane, S. *Science* **256**, 1327-1331 (1992).
10. Desimone, R. *J. cogn. Neurosci.* **3**, 1-8 (1991).
11. Zeki, S. & Shipp, S. *Nature* **335**, 311-317 (1988).
12. Young, M. P. *Nature* **358**, 152-155 (1992).
13. Ungerleider, L. G. & Mishkin, M. in *Analysis of Visual Behavior* (eds Ingle, D. J., Goodale, M. A. & Mansfield, R. J. W.) 549-586 (MIT Press, Cambridge, Massachusetts, 1982).
14. Maunsell, J. H. R. & Newsome, W. T. *A. Rev. Neurosci.* **10**, 363-401 (1987).
15. Eccles, J. C. *Proc. R. Soc. B* **227**, 411-428 (1986).
16. Barbas, H. in *Advances in Neurology* Vol. 57 (eds Chauvel, P., Delgado-Escueta, A. V., Halgren, E. & Bancaud, J.) 91-115 (Raven, New York, 1992).
17. Fuster, J. M. *The Prefrontal Cortex*, 2nd edn (Raven, New York, 1989).
18. Fuster, J. M. *Curr. Opin. Neurobiol.* **3**, 160-165 (1993).
19. Gerfen, C. R. in *Motor and cognitive functions of the prefrontal cortex* (eds Thierry, A.-M., Glowinski, J., Goldman-Rakic, P. S. & Christen, Y.) 78-92 (Springer, Berlin, 1994).
20. Barbaumer, N., Elbert, T., Canavan, A. G. M. & Rockstroh, B. *Physiol. Rev.* **70**, 1-41 (1990).
21. Saint-Cyr, J. A., Ungerleider, L. G. & Desimone, R. *J. comp. Neurol.* **298**, 129-156 (1990).
22. Sherik, J. in *Cerebral Cortex* Vol. 5 (eds Jones, E. G. & Peters, A.) 467-499 (Plenum, New York, 1986).

23. Fries, W. *Vis. Neurosci.* **4**, 205-216 (1990).
24. Sparks, D. L. *Physiol. Rev.* **66**, 118-171 (1986).
25. Ungerleider, L. G., Desimone, R., Galkin, T. W. & Mishkin, M. *J. comp. Neurol.* **223**, 368-386 (1984).
26. Robinson, D. L. & Petersen, S. E. *Trends Neurosci.* **15**, 127-132 (1992).
27. Cowey, A. & Stoerig, P. *Trends Neurosci.* **14**, 140-145 (1991).
28. Cowey, A. & Stoerig, P. *Nature* **373**, 247-249 (1995).
29. Barbur, J. L., Watson, J. D. G., Frackowiak, R. S. J. & Zeki, S. *Brain* **116**, 1293-1302 (1993).
30. Blake, R. *Psychol. Rev.* **96**, 145-167 (1989).
31. Myerson, J., Miezin, F. & Allman, J. *Behav. Analysis Lett.* **1**, 149-159 (1981).
32. Logothetis, N. K. & Schall, J. D. *Science* **245**, 761-763 (1989).
33. Crick, F. & Jones, E. *Nature* **361**, 109-110 (1993).
34. Land, E. H. & McCann, J. J. *J. opt. Soc. Am.* **61**, 1-11 (1971).
35. Blackwell, S. T. & Buchsbaum, G. *J. opt. Soc. Am.* **A5**, 1772-1780 (1988).
36. Zeki, S. *Nature* **282**, 412-418 (1980).
37. Zeki, S. *Proc. R. Soc. B* **217**, 449-470 (1983).
38. Schein, S. J. & Desimone, R. *J. Neurosci.* **10**, 3369-3389 (1990).
39. DeValois, R., Albrecht, D. G. & Thorell, L. G. *Vision Res.* **22**, 545-559 (1982).
40. Levitt, J. B., Kiper, D. C. & Movshon, J. A. *J. Neurophysiol.* **71**, 2517-2542 (1994).
41. Pickersgill, M. J. Q. *Jl exp. Psychol.* **11**, 168-172 (1961).
42. Roland, P. E. & Gulyas, B., *inter alia Trends Neurosci.* **17**, 281-297 (1994).
43. Crick, F. *The Astonishing Hypothesis* (Scribner's, New York, 1994).

ACKNOWLEDGEMENTS. F.C. is supported by the J. W. Kieckhefer Foundation. C.K. is supported by the Air Force Office of Scientific Research, the Office of Naval Research and the National Science Foundation. We thank L. G. Ungerleider, and D. I. MacLeod and S. He for allowing us to quote unpublished experimental work, and the Santa Fe Institute for their hospitality.

A population of very diffuse Lyman- α clouds as the origin of the He⁺ absorption signal in the intergalactic medium

A. Songaila, E. M. Hu & L. L. Cowie

Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, Hawaii 96822, USA

KNOWLEDGE of the physical state of the relatively uniform component of the intergalactic medium (the 'substrate') is critical to understanding the propagation of ionizing radiation and dynamical energy through intergalactic space, and for establishing the boundary conditions for the formation of the intergalactic gas clouds and galaxies that are assumed to have condensed from it. Uniformly distributed hydrogen, and, even more so, He⁺ will produce characteristic smooth absorption in the spectra of high-redshift quasars¹⁻⁴, but at low spectral resolution it is difficult to distinguish such an absorption trough from the cumulative effect of absorption by the Lyman- α 'forest' of clouds. We report the detection of a population of weak 'forest' clouds with column density down to $2 \times 10^{12} \text{ cm}^{-2}$, and show that absorption in these clouds can account for a recent measurement¹ of strong He⁺ absorption without necessarily having to invoke a diffuse intergalactic medium.

The Lyman- α opacity of any uniformly distributed neutral hydrogen in intergalactic space between ourselves and distant quasars would depress the quasar continuum on the short-wavelength side of Lyman- α emission, producing a uniform absorption trough, the so-called 'Gunn-Peterson effect'², but only an upper limit on this effect is observed^{3,4}. The ratio He⁺/He²⁺ should be larger than H⁰/H⁺, motivating searches for the corresponding Lyman- α absorption trough from He⁺. Such a detec-

TABLE 1 Ionization balance at redshift $z = 3.2$

Observed z	C ³⁺ /Si ³⁺	Si ⁺ /Si ³⁺	$\log_{10} \Gamma$	S_L
3.3364	5.5	0.04*	-1.2 (-0.9)	95 (126)
3.3322	1.8	0.45	-2.7 (-2.0)	74 (84)
3.1908	3.4	2.4	-3.2 (-2.6)	21 (32)
3.1719	2.0	0.51	-2.7 (-2.0)	63 (74)

Both $\log_{10} \Gamma$ and S_L (defined in the text) have been computed for an optically thin cloud with $N(\text{H I}) \ll 10^{17} \text{ cm}^{-2}$ and (in parentheses) with $N(\text{H I}) = 10^{19} \text{ cm}^{-2}$. S_L is extremely insensitive to the cloud optical depth.

* Inferred from C⁺/Si³⁺.

tion of strong He⁺ absorption has recently been claimed¹ in the quasar Q0302-003 below the quasar's redshift $z = 3.286$, with optical depth $\tau > 1.7$ at the 90% confidence level. As the authors of ref. 1 fully recognized, it is unfortunately extremely difficult to distinguish a true contribution from the intergalactic medium (IGM) at the spectral resolution (~ 100) of their Hubble Space Telescope (HST) Faint Object Camera observations; they had to proceed by modelling the contribution to τ from known Lyman- α clouds and concluded that it was unlikely to be as large as the observed value. However, as they and others⁵ have pointed out, this conclusion depends crucially on knowing the number density of low-column-density neutral hydrogen clouds in the Lyman forest, until now a very poorly determined quantity. We report here observations of the Lyman forest in Q0302-003 at high spectral resolution (36,000) and very high signal-to-noise (45 per resolution element) using the high resolution HIRES echelle spectrograph⁶ on the Keck 10-m telescope on Mauna Kea, Hawaii. With this improvement in resolution over previous⁷⁻⁹ data, we show that the Lyman-forest column density spectrum rises steeply at the low end—that is, there are many low-column-density clouds. This removes one of the major uncertainties and strengthens the likelihood that the He⁺ absorption is produced by the forest. Whether the observed Lyman- α clouds are therm-

FIG. 1 Top, high-resolution spectrum of the quasar Q0302-003 (emission redshift $z_{\text{em}} = 3.286$), which is the sum of eight 2,400-s exposures made with the HIRES spectrograph at the W. M. Keck 10-m telescope on Mauna Kea, Hawaii, on the nights of 1994 September 16 and October 14 UT. The data were reduced as described in ref. 15; the resulting spectrum has a resolution of 36,000 and a signal-to-noise ratio of 45 per resolution element. The portion of the spectrum between 4,000 and 5,500 Å is shown. At wavelengths $> 5,100$ Å, the CCD (charge-coupled device) does not fully cover the echelle orders, and inter-order gaps are present in the spectrum. The sharp features seen at wavelengths longer than the quasar Lyman- α emission are intergalactic metal-line absorption. At shorter wavelengths there is an extremely rich Lyman forest. It has been suggested⁸ that the apparent gap in the forest at 5,050-5,100 Å may arise from enhanced ionization by a foreground quasar (the 'proximity effect'). Bottom, the 400 Å segment of the normalized spectrum in the Lyman forest immediately to the blue of the quasar Lyman- α emission wavelength.

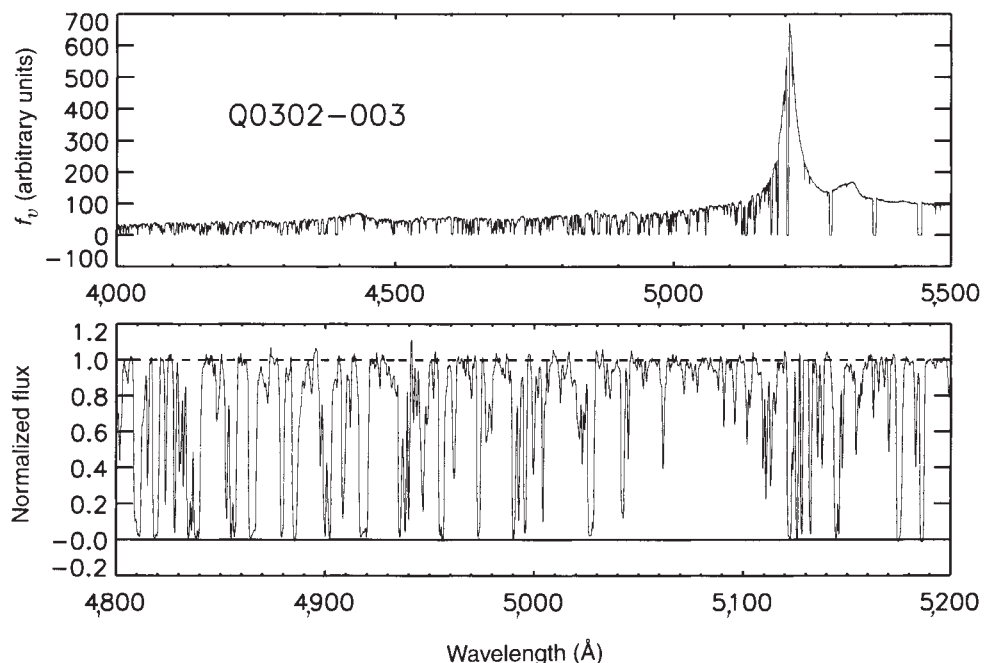
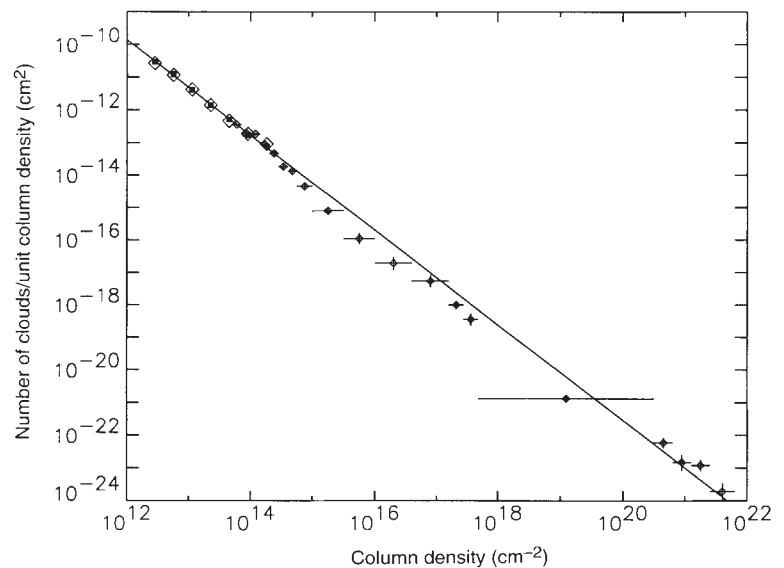


FIG. 2 Number density of Lyman-forest clouds per unit column density and redshift as a function of column density. Intermediate-column-density data as compiled by Petitjean *et al.*¹⁶ is shown as small open diamonds, overlaid with the associated error bars. The low-column-density data of Hu *et al.*¹¹ is shown as filled squares (average of the whole sample), with the Q0302–003 subsample overlaid as large open diamonds. The straight line is the best power-law fit in the region (2×10^{12}) – (3×10^{14}) cm^{-2} , with index -1.46 .



ally or turbulently broadened is also important, because if the hydrogen is thermally broadened, then the He absorption is half as wide, decreasing the average He opacity. Recent observations¹⁰ of weak metal lines corresponding to Lyman-forest clouds imply that turbulent broadening is indeed the dominant mechanism, with each Lyman- α cloud being not a single thermally broadened line but a blend of narrower components. We shall quantify this below.

Our spectrum of Q0302–003 covers the wavelength range 3,600–6,000 Å; the region between 4,000 and 5,500 Å is shown in Fig. 1 (top). The blaze function of the spectrograph in each order was removed by normalizing to fits to white-dwarf standard stars obtained in the same configuration before and after each quasar observation. The continuum level of the spectrum was then established by an iterative fit to each spectral order: all points above a critical value were fitted by a third-order polynomial and renormalized; the critical value was then raised until the extrema in the noise in the Lyman-forest region were matched to the extremal noise characteristics in the region above the rest-frame Lyman- α emission (taking into account the system response determined from the white-dwarf observations and the overall spectral shape of the quasar). The normalized spectra

were then spliced together to form a total spectrum. Figure 1 (bottom) shows the normalized spectrum between 4,800 and 5,200 Å in the forest. This continuum-fitting procedure assumes that there are regions of the spectrum which are free of opacity—that is, they have no weak hydrogen clouds or source of uniform opacity. If this assumption is wrong, as it may be, the continuum should be at a higher value, and the net opacity of both hydrogen and He⁺ will be larger than our derived values. With the adopted continuum, the mean optical depth caused by the Lyman forest is $\langle \tau(\text{Ly}\alpha) \rangle = 0.31$ between 4,500 and 5,000 Å. More importantly, $\sim 75\%$ of this spectral region has an optical depth $\tau(\text{Ly}\alpha) > 0.05$ (Fig. 1). Much of our detailed discussion follows immediately from this. In particular, if it is assumed that hydrogen and helium absorption lines have the same width (turbulent broadening) and that $\tau(\text{He II})/\tau(\text{H I}) = 20$, all of these clouds will have saturated helium absorption, and the average optical depth in He⁺ from the forest clouds will exceed $\tau(\text{He}^+) = 1.4$ immediately, without having to consider any weaker clouds with $\tau(\text{Ly}\alpha) < 0.05$.

This structure of the optical depths implies that there are many low-column-density neutral hydrogen clouds in the intergalactic medium. A full analysis is complicated by the problems of cloud

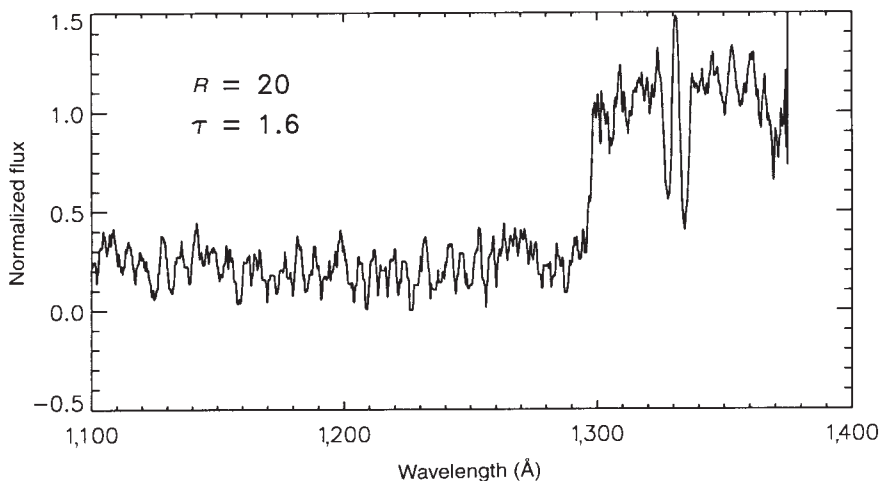


FIG. 3 Simulated He II absorption spectrum for $R \equiv \tau(\text{He II})/\tau(\text{H I}) = 20$. The mean opacity is $\tau = 1.6$. For higher values of $\tau(\text{He II})/\tau(\text{H I})$, the simulated spectrum is consistent with the observed HST spectrum.

blending and will be given elsewhere¹¹, but we summarize the results in Fig. 2, which shows the number density of forest clouds as a function of column density to $N(\text{H I}) = 2 \times 10^{12} \text{ cm}^{-2}$. The best-fit power law between 2×10^{12} and $2 \times 10^{14} \text{ cm}^{-2}$ has an index of -1.5 , and it is the steep rise in the number of clouds at low column density that results in the weak optical depth structure in the spectrum.

But because of the very high resolution of these observations (which effectively resolve the Lyman-forest clouds), it is not actually necessary to perform an analysis of the cloud spectrum to predict the He^+ absorption. Rather, we can use the trend of optical depth with wavelength seen in hydrogen to predict that of He^+ . This analogue method removes many of the uncertainties in the modelling. In particular, if we assume that the absorption lines of singly-ionized helium and hydrogen have the same shape (equivalent to the assumption of turbulent broadening) and that the optical depth of He^+ is in a fixed ratio R to that of H^0 , then the spectrum of the helium forest is

$$S(\lambda) = \exp[-R\tau(4\lambda)] = [S(4\lambda)]^R \quad (1)$$

where S is the normalized spectrum and τ is the optical depth at wavelength λ . We have constructed the He^+ spectrum using this method, smoothed it to the resolution of the HST observations, and measured the optical depth across the He^+ break. We find that, for $R=20$ (illustrated in Fig. 3), the measured mean He^+ optical depth is $\langle\tau(\text{He}^+)\rangle = 1.6$.

As Miralda-Escudé¹² has pointed out, if the spectrum of cloud column densities is a power law with index β extending to column densities lower than those at which the He^+ optical depth in a line is near unity (that is, $N(\text{H I}) \approx 2 \times 10^{12} \text{ cm}^{-2}$ for $R=20$) then $\langle\tau(\text{He}^+)\rangle$ scales as

$$\langle\tau(\text{He}^+)\rangle = 1.6 \left(\frac{R}{20}\right)^{\beta-1} \xi^{2-\beta} = 1.6 \left(\frac{R\xi}{20}\right)^{0.5} \quad (2)$$

where ξ is the ratio of the He^+ line width to that of hydrogen. ($\xi=0.5$ for thermal broadening and $\xi=1$ for turbulent broadening.) This scaling law allows our result to be generally extended.

The recent detection of weak lines corresponding to metals in the forest clouds permits a determination of the ξ parameter. Cowie *et al.*¹⁰ find a mixture of thermal and turbulent broadening in roughly comparable amounts, giving $\xi \approx 0.8$. This then implies that we require $R > 28$ to account for the $\tau > 1.7$ (90% confidence limit) of Jakobsen *et al.*¹.

The quantity R is determined by the shape of the ionizing ultraviolet spectrum, and is equivalent to $0.43 J(\text{H I})/J(\text{He II})$, where $J(\text{H I})/J(\text{He II})$ is the ratio of the ionizing flux at the H I ionization edge to that at the He II edge^{5,12}; the condition on R then becomes $S_L \equiv J(\text{H I})/J(\text{He II}) > 65$. It is possible to measure S_L using observations of relative ionization stages of IGM metal lines¹³. In particular, because the ionization potential of C^{3+} is very close to that of He^+ , the ratio of C^{3+} to lower ionization stages is extremely sensitive to S_L . The methodology (A.S. and L.L.C., manuscript in preparation) is to measure the ionization parameter Γ (the ratio of the number density of hydrogen ionizing photons to the particle density), using species with ionization energies lower than that of He^+ (here Si^+ and Si^{3+}), and then to infer the amount of C^{3+} as a function of Γ and S_L . To carry out this measurement at $z=3.2$ we have observed the quasar Q2000-330 (emission redshift 3.77) with HIRES and obtained ionization ratios for all the C^{3+} lines with $N(\text{C}^{3+}) > 8 \times 10^{12} \text{ cm}^{-2}$ between the redshifts of 3.0 and 3.4. These measurements are summarized in Table 1. To estimate Γ we used the ratio $\text{Si}^+/\text{Si}^{3+}$ assuming a Bechtold spectrum¹⁴. We considered cases where these clouds had $N(\text{H I}) \leq 10^{17} \text{ cm}^{-2}$ (optically thin clouds) and $N(\text{H I}) = 10^{19} \text{ cm}^{-2}$. With Γ determined, S_L was then measured from the ratio $\text{C}^{3+}/\text{Si}^{3+}$. The measured S_L (listed in Table 1) are highly insensitive to the assumptions, and we find a median S_L of 70 at $z=3.2$. We can see from this that the ionization field in the IGM at $z=3.2$ is in fact sufficiently soft that the He^+ forest can account for the Jakobsen *et al.*¹ result.

We conclude that the He^+ trough in the spectrum of Q0302-003 may be produced by Lyman- α clouds, and that this opacity in turn can produce a self-consistently soft intergalactic ultraviolet flux at these redshifts. We emphasize that this does not rule out the possibility of a true He^+ Gunn-Peterson effect at $z=3.2$, but we feel that the current observations do not require it. $\square$

Received 9 November 1994; accepted 27 March 1995.

1. Jakobsen, P. *et al.* *Nature* **370**, 35-39 (1994).
2. Gunn, J. E. & Peterson, B. A. *Astrophys. J.* **142**, 1633-1636 (1965).
3. Jenkins, E. B. & Ostriker, J. P. *Astrophys. J.* **376**, 33-42 (1991).
4. Webb, J. K., Barcons, X., Carswell, R. F. & Parnell, H. C. *Mon. Not. R. astr. Soc.* **255**, 319-324 (1991).
5. Madau, P. & Meiksin, A. *Astrophys. J.* **433**, L53-L56 (1994).
6. Vogt, S. S. *et al.* *S.P.I.E.* **2198**, 362-375 (1994).
7. Sargent, W. L. W., Steidel, C. C. & Boksenberg, A. *Astrophys. J. Suppl. Ser.* **69**, 703-761 (1989).
8. Dobrzycki, A. & Bechtold, J. *Astrophys. J.* **377**, L69-L72 (1991).
9. Bechtold, J. *Astrophys. J. Suppl. Ser.* **91**, 1-78 (1994).
10. Cowie, L. L., Songaila, A., Kim, T.-S. & Hu, E. M. *Astr. J.* (in the press).
11. Hu, E. M., Kim, T.-S., Cowie, L. L., Songaila, A. & Rauch, M. *Astr. J.* (submitted).
12. Miralda-Escudé, J. *Mon. Not. R. astr. Soc.* **262**, 273-276 (1993).
13. Steidel, C. & Sargent, W. L. W. *Astrophys. J.* **343**, L33-L35 (1989).
14. Bechtold, J., Weymann, R. J., Lin, Z. & Malkan, M. A. *Astrophys. J.* **315**, 180-197 (1987).
15. Songaila, A., Cowie, L. L., Hogan, C. J. & Rugers, M. *Nature* **368**, 599-604 (1994).
16. Petitjean, P., Webb, J. K., Rauch, M., Carswell, R. F. & Lanzetta, K. *Mon. Not. R. astr. Soc.* **262**, 499-505 (1993).

ACKNOWLEDGEMENTS. We are grateful to J. Aycok and T. Bida for help with obtaining the observations, and to P. Jakobsen, J. Miralda-Escudé and M. Shull for their helpful comments on an earlier draft of this paper. The authors were visiting astronomers at the W. M. Keck Observatory, jointly operated by the California Institute of Technology and the University of California. This work was supported by the State of Hawaii and by NASA.

Superconductivity and cation-vacancy ordering in the rare-earth fulleride $\text{Yb}_{2.75}\text{C}_{60}$

E. Özdaş^{*†}, A. R. Kortan^{*}, N. Kopylov^{*},
A. P. Ramirez^{*}, T. Siegrist^{*}, K. M. Rabe[†],
H. E. Bair^{*}, S. Schuppler^{*†} & P. H. Citrin^{*}

^{*} AT&T Bell Laboratories, Murray Hill, New Jersey 07974, USA

[†] Department of Applied Physics, Yale University, New Haven, Connecticut 06520, USA

INTERCALATED fullerides exhibit remarkable physical properties, including superconductivity at record high temperatures for organic compounds¹⁻⁶. The donor intercalants have so far been limited to alkali and alkaline-earth metals. Here we describe the synthesis, structure and magnetic properties of a rare-earth-doped fulleride, $\text{Yb}_{2.75}\text{C}_{60}$, which exhibits superconductivity below 6 K. Ytterbium cations occupy off-centred interstitial sites in the face-centred cubic C_{60} lattice, and leave eight tetrahedral sites vacant in an orthorhombic unit cell which is thereby doubled in each direction. These cation-vacancy sites exhibit long-range ordering, generating a superstructure which seems to be stabilized by a short-range interaction between charge-deficient single bonds in C_{60} and exclusively divalent cations. Our results demonstrate that superconductivity is not limited to C_{60} compounds of high structural symmetry.

We studied the physical properties of the Yb_xC_{60} system in the range $0 < x < 5$. The samples were prepared from high purity Yb (Ames Laboratory, Ames, Iowa) and chromatographed C_{60} (Scientific Exchange Inc. Center, Sandwich, New Hampshire). Weighed powders were mixed and loaded into custom-made

[‡] Present addresses: Hacettepe University, Beytepe, Ankara, Turkey (E.Ö.); Institut für Nukleare Festkörperphysik, Kernforschungszentrum Karlsruhe, D-76021 Karlsruhe, Germany (S.S.).

tungsten cells in a controlled-atmosphere glove box with O₂ and H₂O levels of <1 p.p.m. The sample cells, sealed in quartz tubes under 10⁻⁷ torr, were heated at optimum reaction conditions of 630 °C for 6 h (with a few intermediate grindings), and cooled to room temperature at a rate of 5 °C min⁻¹. Bulk synthesis of Yb_{2.75}C₆₀ could only be achieved in a narrow temperature range: above 670 °C, the structure rapidly decomposed to an amorphous phase, and below 630 °C, Yb diffusion limited the reaction. Figure 1 shows the X-ray powder diffraction data (dots) of this novel phase, and our best-fit results (solid line, described below). The X-ray diffraction pattern of Yb_{2.75}C₆₀ shows many new peaks which are not consistent with a face-centred cubic (f.c.c.) structure and which correspond to no previously known structures of intercalated C₆₀ compounds or possible impurity phases, for example Yb₃C or Yb₂O₃. The new reflections can only be indexed as primitive reflections with doubled unit cell parameters of the original f.c.c. unit cell. Although the ordering of the C₆₀ molecules gives rise to simple cubic reflections in the high-*q* region⁷⁻⁹ (*q* is the momentum transfer), these new reflections of the vacancy-ordered superstructure are located at low *q*, similar to Ca₅C₆₀ (ref. 10).

An iterative series of fits to the X-ray diffraction data, represented by the solid line in Fig. 1, led to a quantitative structural determination of Yb_{2.75}C₆₀. The refined parameters are summarized in Table 1, and a schematic view of the structural model is shown in Fig. 2. Yb cations are located in both tetrahedral (T) and octahedral (O) interstitial sites of the cubic close-packed C₆₀ lattice. All of the O-sites are occupied by single cations, which are significantly displaced off-centre. By contrast, one out of every eight T-sites in the subcell is vacant, and this vacancy alternates between adjacent T-sites in each of the *x*, *y* and *z* directions. The alternation is responsible for the doubling of the unit cell, leading to a total of eight such ordered vacancies. Each T-vacancy is surrounded by four O-site cations, whose off-centre displacements are directed towards these vacancies. For example, the T-vacancy at (1/8, 1/8, 1/8) in Fig. 2, shown as a red corner of a white cube connecting the other local T-sites, has four neighbouring cations which move from the centre of the O-sites (that is, the centre of the green cubes) towards the off-centre coordinates (1/5, 1/5, 1/5), (1/20, 1/20, 1/5), (1/20, 1/5, 1/20) and (1/5, 1/20, 1/20), that is, the corners of the green cubes closest to the vacancy. These large displacements, ~2.4 Å, lead to O-cations having three rather than six first-neighbour C₆₀ molecules.

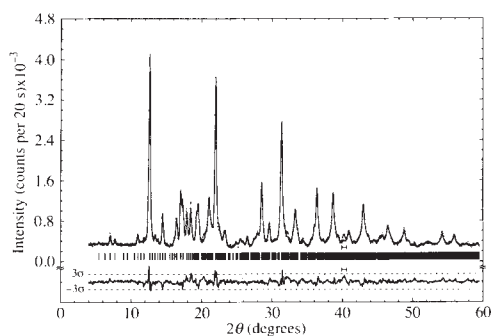


FIG. 1 Powder X-ray diffraction pattern (dots) of Yb_{2.75}C₆₀ phase, collected with Cu K α (1.54 Å) radiation using a 12-kW rotating-anode generator equipped with a triple-axis goniometer and highly oriented ZYA graphite monochromator crystals. Least-squares fit (solid line) used the DBWS Rietveld refinement¹⁸. A total of 3,029 reflections were measured and refined, 407 of which had intensity levels >3 σ . Refinements using 30 parameters yielded $R_{\text{Bragg}}=0.044$, $R_p=0.055$, a weighted *R*-factor of 0.065, and *S* (or χ^2)=1.48. The region contaminated with a W reflection (39.9°–40.6°) was excluded from the refinement. Ticks denote allowed Bragg reflections of superstructure in *Pcab* symmetry, and the lower trace shows the difference (data – model) spectrum.

TABLE 1 Structural parameters of Yb_{2.75}C₆₀ from Rietveld refinements

Atom	Site	x	y	z	B(Å ²)	N
Yb(1 1)	8c	0.1263 (2)*	0.1162 (3)	0.3753	5.0 (1)	1.00
Yb(1 2)	8c	0.3753	0.1263	0.1162	5.0	1.00
Yb(1 3)	8c	0.1162	0.3753	0.1263	5.0	1.00
Yb(2 1)	8c	0.1263 (2)	0.3630 (3)	0.3775	5.0	1.00
Yb(2 2)	8c	0.3775	0.1263	0.3630	5.0	1.00
Yb(2 3)	8c	0.3630	0.3775	0.1263	5.0	1.00
Yb(3)	8c	0.3747	0.3665	0.3753	5.0	1.00
Yb(4 1)	8c	0.2000 (2)	0.2000	0.2000	3.9 (2)	0.85 (2)
Yb(4 2)	8c	0.0500	0.0500	0.2000	3.9	0.85
Yb(4 3)	8c	0.2000	0.0500	0.0500	3.9	0.85
Yb(4 4)	8c	0.0500	0.2000	0.0500	3.9	0.85
Yb(5 1)	8c	0.2000	0.3000	0.3000	3.9	0.15
Yb(5 2)	8c	0.0500	–0.0500	0.3000	3.9	0.15
Yb(5 3)	8c	0.3000	–0.0500	0.0500	3.9	0.30

Space group, *Pcab* (No. 61, option 2); *a*=27.8743 (2), *b*=27.9804 (2), *c*=27.8733 (2) Å; *Z*=32 formula units per unit cell; calculated density, 2.92 g cm⁻³. The shape of C₆₀ molecules was constrained to icosahedral symmetry. The ratio of the two non-equivalent bond lengths *d*_{6,6}/*d*_{5,6} was fixed at 1.372 Å/1.458 Å obtained by single-crystal x-ray diffraction¹⁶ in the ordered phase of C₆₀. Initial coordinates of the 240 independent carbon atoms defining all C₆₀ molecules of the superstructure were reproduced for the cage diameter of 7.057 Å from values in ref. 16. Occupancy *N* of carbon atoms was fixed as 1, and isotropic temperature factor *B* of carbon was found to be 0.86 (0) Å². (Under 'Site', 8 is the multiplicity and *c* is the Wyckoff letter.)

* Estimated standard deviations of the last digit in parentheses are strictly from refinement.

The cations occupying the T-sites also experience off-centre displacements (albeit much smaller) towards the middle of two neighbouring C₆₀ molecules, that is, along one of the principal *x*, *y* or *z* directions. Refinements testing all three directions for every T-cation converged rapidly to an essentially uniaxial distortion of ~0.3 Å from the centre of the T-site, thus lowering the first-neighbour coordination from four to two C₆₀ molecules. Overall, the *Pa* $\bar{3}$ cubic space group is consistent with the salient features of the model, namely, the T-vacancy ordering, the O-cation displacement toward the T-vacancies, and the displacement of most T-cations (Yb(1 *i*), Yb(2 *i*), *i*=1, 2, 3) in any direction. The predominantly uniaxial displacement of the Yb(3) cation at (3/8, 3/8, 3/8), however, makes *x*, *y* and *z* inequivalent, resulting in a ~0.1 Å contraction of the unit cell along the *x* and *z* directions. The cubic unit cell becomes orthorhombic with the space group *Pcab* (no. 61, option 2), a direct subgroup of *Pa* $\bar{3}$.

Further refinement is obtained by rotating all C₆₀ molecules about one of their local symmetry axes in the *Pcab* space group. We first selected the set proposed¹¹ for the 2*a* structure of C₆₀ (ref. 12) and obtained the best goodness-of-fit value, *S*=1.65, for a rotation angle near 38°. This is very different from the 97.5° rotation observed in all *Pa* $\bar{3}$ phases of C₆₀ (ref. 7), Na₂C₆₀ (ref. 8) and Na₂CsC₆₀ (ref. 9) compounds, where electron-rich double bonds face five-membered rings of neighbouring molecules and alkali cations. In Yb_{2.75}C₆₀, the orientational ordering of C₆₀ molecules aligns electron-poor five-membered rings with Yb cations. As a total of 56% of all five-membered rings in all molecules face a cation position in the above fits, we searched for a set of special axes which further optimized occupation of such five-membered ring configurations. In Yb_{2.75}C₆₀, there are three different types of C₆₀ molecules (Fig. 2*b*: C₆₀(1), C₆₀(2 *i*) (*i*=1, 2, 3), and C₆₀(3)), with two, one and zero neighbouring T-vacancies, respectively. The maximum occupation of five-membered rings occurs when 28 of the total 32 molecules (4 C₆₀(1) and 24 C₆₀(2 *i*)) rotate about the axis passing through the vacancies. The remaining C₆₀(3) molecules have similar ordering with T-cations, but only when rotated about their local {111} axes. Taking into account all of the axes mentioned with 37.5° rotations, we obtain a total of 75% occupation of all five-membered rings and an even lower *S* value of 1.48. Furthermore, all C₆₀ molecules are themselves oriented

with respect to one another, positioning their double bonds over either a six-membered ring or double bond in a neighbouring molecule. Other sets of rotations, including the ones around two-fold axes, do not orient nearly as many five-membered rings with the cation positions.

Figure 3 shows the temperature dependence of d.c. magnetization and differential microwave absorption for $\text{Yb}_{2.75}\text{C}_{60}$. Microwave loss measurements on powder samples indicate a field-dependent absorption near 6 K. Below this temperature, the sample is superconducting. The sample was zero-field cooled (ZFC) to 2 K, then a field of 5 G was applied and the magnetization of the shielding branch was measured as a function of temperature in the range 2–10 K. After cooling from above the superconducting transition temperature, T_c , in the applied field, the magnetic flux expulsion was measured in the superconducting state (the Meissner branch). The Meissner fraction is determined to be 8% of the ZFC value, which is 5.6% volume fraction (see Fig. 3). This suggests that in such a low field, the flux expulsion is incomplete, and most of the magnetic flux is trapped and pinned during the cooling process owing to weak intergranular coupling.

Two independent photoemission studies^{13,14} on Yb– C_{60} film samples have suggested possible mixed-valent character (that is, both Yb^{2+} and Yb^{3+}) with Yb_2C_{60} stoichiometry¹³. The possibility of multiple configurations in Yb_xC_{60} was investigated

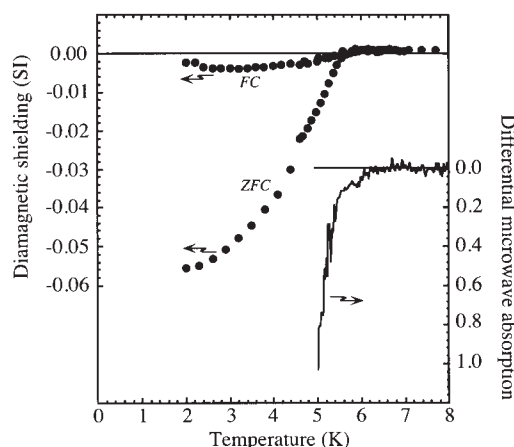


FIG. 3 Superconductivity in $\text{Yb}_{2.75}\text{C}_{60}$, as seen by magnetic susceptibility and microwave absorption measurements. SQUID data are shown for both field-cooled (FC) and zero-field cooled (ZFC) cases.

using near-edge and extended X-ray absorption fine structure (NEXAFS and EXAFS) measurements (S.S. *et al.*, manuscript in preparation). The NEXAFS data gave clear evidence for only divalent Yb in the bulk material. The EXAFS data showed Yb in two well defined environments with average first-neighbour Yb–C distances differing by 0.18 ± 0.03 Å, and relative populations of Yb in the two environments of 1.7 ± 0.3 . These results are consistent with Yb^{2+} cations occupying T- and O-sites in a single configuration. Combined with our finding here of $x = \text{Yb(T)} + \text{Yb(O)} = 2.75$, they imply $\text{Yb(T)} \approx 1.7$, and $\text{Yb(O)} \approx 1$, in excellent agreement with our detailed structural determination.

The underlying physical origin of this strikingly different ordering is due not to cation size—the ionic radii of Yb^{2+} and Na^+ are nearly identical—but is instead attributed to a strong, short-range, directional interaction involving the divalent Yb cations. The most dramatic evidence for this is found in the T-vacancies and the associated off-centre displacements of the O-cations. These large displacements go beyond lowering the first-neighbour coordination of C_{60} molecules from six to three; the O-cations are even displaced with respect to the centre of the five-membered ring towards two of the five C atoms. Analogous displacements are also found for T-cations, whose first-neighbour coordination is reduced from four to two C_{60} molecules. This strong directional ordering is manifest in average nearest-neighbour Yb–C distances of 2.6 and 2.8 Å for the T- and O-sites, respectively. Although these Yb–C distances are ~ 0.2 – 0.4 Å longer than typical σ bonds in, for example, Yb carbide¹⁵, they are ~ 0.1 and ~ 2.5 Å shorter than those for the respective T- and O-cations in the $\text{Pa}\bar{3}$ structure of, for example, $\text{Na}_2\text{RbC}_{60}$ (ref. 16). Such $\text{Pa}\bar{3}$ structures containing weak electrostatic (ionic) interactions lead to an order-disorder transition around 300 K (refs 16, 17). Differential scanning calorimeter measurements on $\text{Yb}_{2.75}\text{C}_{60}$ samples, however, showed no signs of any phase transition up to 873 K. The collective evidence, therefore, clearly indicates a strong, short-range, directional interaction between Yb and C_{60} , consistent with the description of predominantly covalent bonding between Yb and C. □

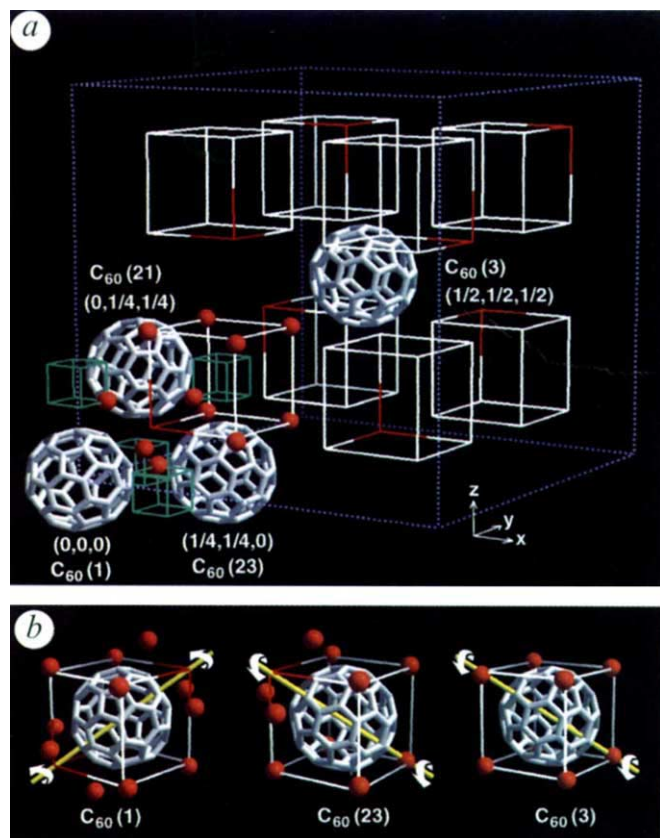


FIG. 2 a, Model of vacancy-ordered structure of $\text{Yb}_{2.75}\text{C}_{60}$, b, three different types of C_{60} molecules with respect to their nearest-neighbour Yb cations. For clarity, only four of the five basis molecules, and eleven basis Yb cations are shown in a. Tetrahedral (T) vacancies are denoted by red corners of white cubes that connect T-sites. The molecules at $(0, 0, 0)$, $(0, 1/4, 1/4)$, $(1/4, 0, 1/4)$, $(1/4, 1/4, 0)$ and $(1/2, 1/2, 1/2)$ were rigidly rotated clockwise 37.5° around $[111]$, $[111][111]$, $[111]$ and $[111]$ axes, respectively. Corners of the green cubes represent eight available sites for the displaced cation in the octahedral site. Red balls indicate Yb cation positions.

Received 20 January; accepted 30 March 1995.

1. Ehrenreich, H & Spoenen, F. (eds) *Solid State Physics* Vol. 48 (Academic, San Diego, 1994).
2. Hebard, A. F. *et al.* *Nature* **350**, 600 (1991).
3. Holczer, K. *et al.* *Science* **252**, 1154 (1991).
4. Rosseinsky, M. J. *et al.* *Phys. Rev. Lett.* **66**, 2830 (1991).
5. Stephens, P. W. *et al.* *Nature* **351**, 632 (1991).
6. Kortan, A. R. *et al.* in *Progress in Fullerene Research* (eds Kuzmany, H., Fink, J., Mehring, M. & Roth, S.) 239–243 (World Scientific, Singapore, 1994).
7. David, W. I. F. *et al.* *Nature* **353**, 147 (1991).
8. Yildirim, T. *et al.* *Phys. Rev.* **B71**, 1383 (1993).

9. Prassides, K. et al. *Science* **263**, 950 (1994).
10. Kortan, A. R. et al. *Nature* **355**, 529 (1992).
11. Harris, A. B., Sachidanandam, R. & Yildirim, T. *Phys. Rev. B* **50**, 2622 (1994).
12. Van Tendeloo, G. et al. *Phys. Rev. Lett.* **69**, 1065 (1992).
13. Ohno, T. R. et al. *Phys. Rev. B* **46**, 10437 (1992).
14. Xia, B., Ruckman, M. W. & Strongin, M. **48**, 14623 (1993).
15. Atoji, M. & Flowers, R. H. J. *chem. Phys.* **52**, 6430 (1970).
16. Burgi, H. B., Restori, R. & Schwarzenbach, D. *Acta crystallogr.* **B49**, 832 (1993).
17. Yildirim, T. et al. in *Progress in Fullerene Research* (eds Kuzmany, H., Fink, J., Mehring, M. & Roth, S.) 235–238 (World Scientific, Singapore, 1994).
18. Wiles, D. B. & Young, R. A. *J. appl. Crystallogr.* **14**, 149 (1981).

ACKNOWLEDGEMENTS. We thank S. Glarum and K. Lyons for valuable discussions. Financial support for E.Ö. from the Scientific and Technical Research Council of Turkey through a NATO Scholarship, and for S. Schuppler from the Alexander-von-Humboldt Foundation, is gratefully acknowledged. The X-ray absorption experiments were performed at the NSLS, Brookhaven National Laboratory, which is supported by the US DOE.

Visualization of fast energy flow and solvent caging in unimolecular dynamics

Paul A. Rejto^{*†}, Eckart Bindewald^{*}
& David Chandler^{*‡}

^{*} Department of Chemistry, University of California, Berkeley, California 94720, USA

ENERGY flow in solution between physically or chemically evolving solute molecules and the surrounding solvent significantly affects the nature of chemical dynamics in liquids. It determines the extent to which the statistical theory of reaction rates^{1,2} is valid; the transfer of energy between solute and solvent influences the ease with which the transition state evolves into the products—the process central to transition-state theory. But analysing the energy flow in liquid-phase dynamics is difficult because these systems are so complex, and the degrees of freedom are consequently so numerous. Here we present a way to address this challenge. We introduce an approach for visualizing the energy flow directly, and apply it to the isomerization of cyclohexane (between boat and chair conformations) in liquid carbon disulphide, a process for which detailed information about the molecular motions is available from molecular dynamics simulations⁸. Our method reveals in pictorial form the formation and relaxation of a solvent cage, and shows that the relaxation has a strong effect on energy flow to and from the transition state on sub-picosecond timescales. We anticipate that this visualization approach will be generally useful for elucidating dynamical molecular processes in solution.

Recent work has demonstrated the existence of fast dielectric relaxation processes that affect charge transfer in highly polar liquids³. We now show that there is an additional class of chemically relevant fluctuations of the solvent which have similarly fast relaxation. In the non-polar case, these fluctuations (or solvent motions) correspond to the reorganization and buffeting of the molecules near the evolving species. The specific mechanism by which energy flows between a solute and solvent, however, is subtle. The constraints of packing forces that characterize the structure of liquids⁴ cause the amplitudes of motions to be small. When atomic trajectories of the solute and its neighbours are visualized with computer graphics, only the impression of apparently random fluttering of atoms is obtained. No reasonably specific pathways for energy flow are revealed.

Here we introduce the concept of 'work landscapes'; use of this approach reveals specific pathways for energy flow between solute and solvent. We chose the isomerization of cyclohexane for the first application of this technique for reasons related to

both experiment and theory. Experiments have shown that the interconversion of the chair and boat isomers of cyclohexane in solution is accelerated by increasing the pressure of the solvent^{5–7}. Transition-state theory would explain this phenomenon if the free-energy barrier to isomerization decreased with increasing solvent pressure. But using reasonably realistic intermolecular potential functions, a variety of theoretical calculations have shown that solvents such as liquid carbon disulphide have no significant effect on the free energy of cyclohexane isomerization^{8,9}. The observed pressure acceleration of the chair–boat interconversion would therefore seem to be a dynamical effect, governed solely by the nature of solvent–solute energy flow. Indeed, phenomenological models have shown that the competition between intra- and inter-molecular energy flow could provide an explanation of the experimental observations^{10,11}.

Molecules that successfully undergo a transition between the chair and boat configurations must dissipate energy after crossing the transition state. If too much excess energy remains in the reaction coordinate for too long a time, the molecule may recross the transition state and thereby fail to complete its transition. Some of the energy is transferred from the reaction coordinate to other internal degrees of freedom of the solute. The remaining energy is transferred to the solvent, and this process can be monitored^{12,13} by examining the differential work done by the solvent on the reaction coordinate. The rate at which this work is done is a function of the reaction coordinate phase-space point, θ , and its time derivative, $\dot{\theta}$, as well as all the other coordinates and velocities characterizing a trajectory of the system. The average rate of work done between time t_1 and t_2 for a given θ and $\dot{\theta}$ is

$$\dot{W}(\theta, \dot{\theta}; t_1, t_2) = -(t_2 - t_1)^{-1} \int_{t_1}^{t_2} \dot{\theta} \langle \partial V / \partial \theta \rangle dt \quad (1)$$

where V refers to total potential energy of interaction between solute and solvent.

The angle-bracketed term in equation (1) is an average over an ensemble of trajectories that pass over the transition state at time $t=0$, that is $\theta(0) = \theta_{TS}$. This averaging removes explicit

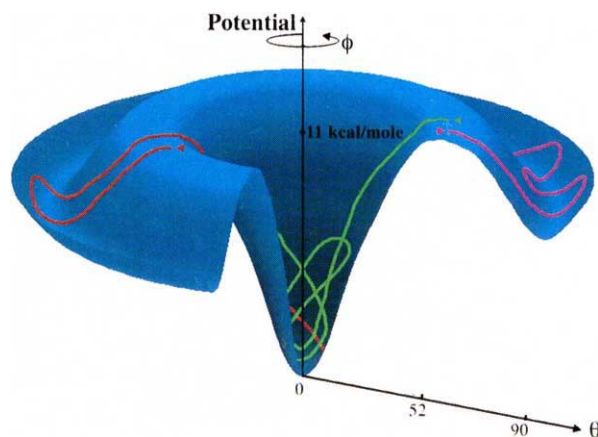


FIG. 1 Representative trajectories projected on to the θ – ϕ potential surface for cyclohexane. The total potential energy and concomitant forces that govern the dynamics are functions of all the atomic coordinates of the system. The blue surface is the potential energy averaged over all coordinates except the collective hindered rotation variables θ and ϕ . The chair configurations of cyclohexane are in the vicinity of $\theta = 0$. The boat configurations are near $\theta = 90^\circ$. The transition-state surface is at $\theta = \theta_{TS} = 52^\circ$. The pink and green trajectories equilibrate without recrossing the transition state in the boat and chair configurations, respectively. These statistical or non-recrossing trajectories are consistent with transition-state theory. The red trajectory on the left is not of this type as it recrosses the transition state before equilibrating in the chair configuration.

[†] Present address: Agouron Pharmaceuticals, 3565 General Atomics Court, San Diego, California 92121, USA.

[‡] To whom correspondence should be addressed.

dependence on other coordinates and velocities, and only their statistical influence is retained. The experimental relevance of this average rate of work done depends on the probability density, $P(\theta, \dot{\theta}; t_1, t_2)$, for observing that particular phase-space point in the time interval $t_1 < t < t_2$. With modern computer graphics, the energy transfer may be visualized by plotting W and P in a single figure which we will call a work landscape.

To draw a work landscape, the pertinent potential-energy topography must be understood well enough to identify meaningful reaction-path variables. For cyclohexane, C_6H_{12} , the hindered rotation coordinate¹⁴ serves as such a variable. Figure 1 shows the potential-energy surface associated with this variable, θ , along with a secondary cyclic generalized coordinate, ϕ (see below). The deep central well corresponds to chair configurations of cyclohexane; the higher-energy outer region of stability corresponds to boat configurations. Owing to the cyclic symmetry of the C_6H_{12} molecule, there are several degenerate chair and boat configurations, and the variable ϕ takes the molecule from one such configuration to another. Both θ and ϕ depend^{14,15} on the positions of all six carbon atoms in cyclohexane. The cartesian coordinate system describing configurations of these atoms is 18-dimensional. The potential pictured in Fig. 1 determines forces in this space. The surrounding solvent atoms give rise to additional forces and cartesian dimensions.

For some systems, the information required to construct work landscapes might be available from experiments, using the technique of frequency-resolved transient absorption. Here we obtained the information from simulations of trajectories that were performed⁸ to calculate thermal rate constants of chair-boat interconversion in carbon disulphide solvent. In computing rate constants, trajectories are initiated at the transition state, with initial phase-space points for all but the reaction coordinate taken from the equilibrium canonical distribution¹⁶⁻¹⁹. We have partitioned these trajectories into two sub-ensembles and thereby create two sets of work landscapes. One sub-ensemble contains those trajectories that do not recross the transition state between chair and boat configurations (as illustrated with two of the three coloured lines in Fig. 1). The other sub-ensemble contains those that do recross (as illustrated by the third coloured line in Fig. 1). Transition-state theory neglects the possibility of recrossing^{1,2,19-21}; it is therefore these recrossing trajectories that are responsible for errors in transition-state theory.

The two types of work landscapes—those coinciding with transition-state theory (non-recrossing trajectories) and those coinciding with its corrections (recrossing trajectories)—are shown in Fig. 2. Three successive time intervals are chosen to illustrate the initial relaxation and the evolution towards equilibration. In the last pair of panels, referring to the time period 0.3–0.6 ps, cyclic patterns have formed. These cycles are not unlike what is expected for the phase-space trajectory of a har-

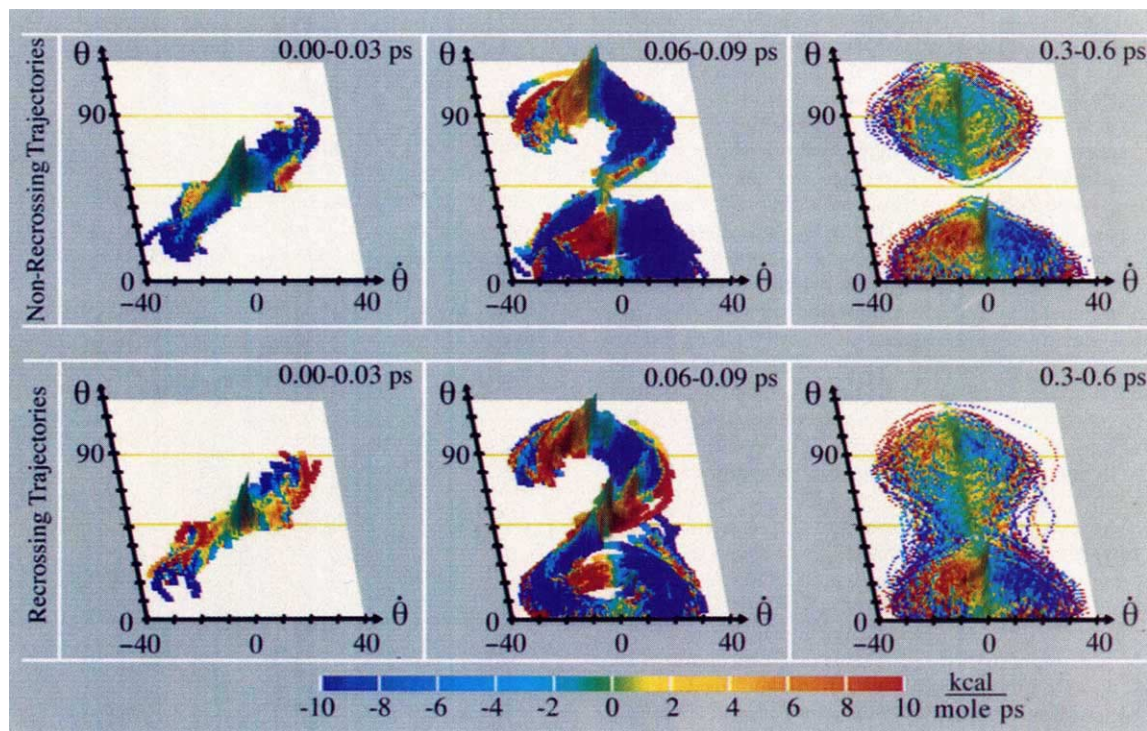


FIG. 2 A series of work landscapes created from 1,000 molecular-dynamics trajectories of cyclohexane isomerizing in liquid carbon disulphide⁸. The initial conditions for the trajectories are taken from the ensemble where the liquid (density 1.5 gm cm^{-3}) and the solute are at room temperature and equilibrated with θ constrained to its transition-state value $\theta_{TS} = 52^\circ$. The trajectories are classified according to type, either recrossing (bottom row of panels) or non-recrossing (top row). For each case, averages are performed for three separate time intervals (left-hand, middle and right-hand panels). For each of the six trajectory

sets thus formed, the work done per unit time (equation (1)) is binned according to the phase-space point where it occurs. Here, a binning volume of $1^\circ \times 1 \text{ rad ps}^{-1}$ was employed. Similarly, for each trajectory set, a visitation number, normalized to form a density, $P(\theta, \dot{\theta}; t_1 t_2)$, was determined for the binning volumes. Plotting these functions creates landscapes, with height corresponding to the probability density and colour corresponding to the work or energy-transfer rate (see colour scale at bottom of figure). (Note that θ and $\dot{\theta}$ are plotted in units of degrees and rad ps^{-1} , respectively.)

monic coordinate. We also point out the red–blue pattern in the chair region (near $\theta=0$, $\dot{\theta}=0$), and the more subtle pattern in the boat region (centred at $\theta=90^\circ$, and $\dot{\theta}=0$). As blue indicates energy transfer from solute to the solvent and red indicates the reverse, the alternating colours are characteristic of harmonic solvent potentials. Such a potential yields

$$\dot{W}(\theta, \dot{\theta}) = -\kappa(\theta - \theta_0)\dot{\theta} \quad (2)$$

where κ is the magnitude of the solvent-imposed restoring force, and θ_0 is a reference point, near 0 for the chair state and near 90° for the boat state. The alternating red–blue patterns are easily understood in terms of this equation: a solvent potential like that leading to equation (2) corresponds to a solvent cage²². For both recrossing and non-recrossing trajectories, solvent caging of stable states is therefore evident after a few tenths of a picosecond.

Consider now the behaviour at earliest times (0–0.03 ps; the two panels at the extreme left of Fig. 2). Away from the line $\dot{\theta}=0$, where the work must vanish, the statistical or non-recrossing trajectories (top-left panel) demonstrate primarily energy transfer to the solvent as evidenced by the large blue regions. By contrast, the recrossing trajectories (bottom-left panel) do not demonstrate any net energy transfer to the solvent. As indicated by equation (2), a solvent cage around the transition state will initially cause the solvent to do negative work on the reaction coordinate, and hence result in a blue landscape at early times. In this regime, the statistical trajectories obeying transition-state theory show the presence of a cage near the transition state. A cage does not always form, however, around a solute in its transition state. We see this in the case of recrossing trajectories, where the work landscape shows no evidence of a cage at the earliest times.

This landscape evolution shows that the restoring forces from the cage are evidently not strong enough to deter motion away from the unstable intramolecular barrier. The cage does, however, allow the solvent to cool the molecule and thus aid its stabilization. The mechanism for the pressure acceleration of the rate is thus clear, as higher pressure will surely enhance formation of the solvent cages that cool activated trajectories. For cyclohexane in liquid carbon disulphide, increasing pressure thereby diminishes the number of recrossing trajectories. Such recrossing trajectories can only lower the rate¹⁶, and a decrease in their number implies an increase in the rate.

The work landscapes can provide yet more detail. During the time period represented in the middle two panels of Fig. 2, a most striking feature is the presence of both blue and red in the chair region near $\theta=20^\circ$ and $\dot{\theta}<0$. This arises from trajectories that begin at the transition state with negative velocity, and follow a path that oscillates about $\theta=0$ (the equilibrium chair configuration). The trajectories are primarily blue during the initial motion away from the transition state, and remain blue for positive $\dot{\theta}$. They become red when $\dot{\theta}$ becomes negative again, which corresponds to the first turning point in the trajectory. The change in sign of the energy transfer in this region indicates a significant change in the interaction between the solvent and the reaction coordinate. It is the signature of the formation of a solvent cage centred at the equilibrium chair configuration. This cage is formed rapidly, in less than 0.1 ps. An analogous process occurs in the boat region.

In the 0.06–0.09 ps period depicted in the middle two panels of Fig. 2, there is a relatively low probability density near the transition state for the non-recrossing trajectories compared with the substantial probability density for the recrossing trajectories. The substantial transition-state density in the latter case is due to those trajectories that never leave the immediate vicinity of the transition state. The remaining recrossing trajectories undergo one or more oscillations before returning to the transition-state region. These resonant recrossings can persist only so long as a solvent cage does not form. After formation of a solvent cage, the molecule will deposit energy into the solvent and begin its relaxation to a stable state. □

Received 12 October 1994; accepted 23 March 1995.

1. Wigner, E. *Trans. Faraday Soc.* **34**, 29–41 (1938).
2. Johnston, H. S. *Gas Phase Reaction Rate Theory* (Ronald, New York, 1966).
3. Jimenez, R., Fleming, G. R., Kumar, P. V. & Maroncelli, M. *Nature* **369**, 471–473 (1994).
4. Chandler, D., Weeks, J. D. & Andersen, H. C. *Science* **220**, 787–794 (1983).
5. Hasha, D. L., Eguchi, T. & Jonas, J. J. *chem. Phys.* **75**, 1571–1573 (1981).
6. Hasha, D. L., Eguchi, T. & Jonas, J. J. *Am. chem. Soc.* **104**, 2290–2296 (1982).
7. Campbell, D. M., Mackowiak, M. & Jonas, J. J. *chem. Phys.* **96**, 2717–2723 (1992).
8. Wilson, M. A. & Chandler, D. *Chem. Phys.* **149**, 11–20 (1990).
9. Singer, S. J., Kuharski, R. A. & Chandler, D. *J. phys. Chem.* **90**, 6015–6017 (1986).
10. Borkovec, M., Straub, J. E. & Berne, B. J. *J. chem. Phys.* **85**, 146–149 (1986).
11. Chandler, D. & Kuharski, R. A. *Discuss. Faraday Soc.* **85**, 329–340 (1988).
12. Rejto, P. A. & Chandler, D. *J. phys. Chem.* **98**, 12310–12314 (1994).
13. Gertner, B. J., Whitnell, R. M., Wilson, K. R. & Hynes, J. T. *J. Am. chem. Soc.* **113**, 74–87 (1991).
14. Pickett, H. M. & Strauss, H. L. *J. Am. chem. Soc.* **92**, 7281–7290 (1970).
15. Strauss, H. L. *J. chem. Educ.* **48**, 221–223 (1971).
16. Keck, J. C. *Adv. chem. Phys.* **13**, 85–121 (1967).
17. Anderson, J. B. *J. chem. Phys.* **58**, 4684–4692 (1973).
18. Bennett, C. H. in *Algorithms for Chemical Computation* (ed. Christofferson, R. E.) 63–97 (ACS Symp. Ser. No. 46, American Chemical Society, Washington DC, 1977).
19. Chandler, D. *J. chem. Phys.* **68**, 2959–2970 (1978).
20. Miller, W. H. *Acc. chem. Res.* **9**, 306–312 (1976).
21. Hänggi, P., Talkner, P. & Borkovec, M. *Rev. mod. Phys.* **62**, 251–341 (1990).
22. Patron, F. & Adelman, S. A. *Chem. Phys.* **152**, 121–131 (1991).

ACKNOWLEDGEMENTS. E. B. thanks the University of California for an exchange studentship for the 1993–94 academic year. Graphics were produced at the University's College of Chemistry Graphics Center, supported by the NIH. Additional support was provided by the US NSF.

Observational evidence for chemical ozone depletion over the Arctic in winter 1991–92

Peter von der Gathen^{*†}, Markus Rex^{*†},
Neil R. P. Harris^{‡§}, Diana Lucic[§],
Bjørn M. Knudsen^{||}, Geir O. Braathen[¶],
Hugo De Backer^{*}, Rolf Fabian[†], Hans Fast[☆],
Manuel Gil^{**}, Esko Kyrö^{††},
Ib Steen Mikkelsen^{||}, Markku Rummukainen^{††},
Johannes Stähelin^{‡‡} & Costas Varotsos^{§§}

^{*} Alfred Wegener Institute for Polar and Marine Research, PO Box 60 01 49, D-14401 Potsdam, Germany

[†] Institute of Environmental Physics, University of Bremen, PO Box 33 04 40, D-28334 Bremen, Germany

[‡] European Ozone Research Coordinating Unit, British Antarctic Survey, Madingley Road, Cambridge CB3 0ET, UK

[§] Centre for Atmospheric Science, Department of Chemistry, University of Cambridge CB2 1EW, UK

^{||} Danish Meteorological Institute, Lyngbyvej 100, DK-2100 Copenhagen Ø, Denmark

[¶] NILU, PO Box 100, Instituttveien 18, N-2007 Kjeller, Norway

^{*} Royal Meteorological Institute, Ringlaan 3, B-1180 Brussel, Belgium

[☆] Atmospheric Environment Service, 4905 Dufferin Street, North York, Ontario M3H 5T4, Canada

^{**} Instituto Nacional de Técnica Aeroespacial, Torrejón de Argoz, 28850 Madrid, Spain

^{††} Finnish Meteorological Institute, Ilmala, SF-99600 Sodankylä, Finland

^{‡‡} Laboratory for Atmospheric Physics, ETH-Hönggerberg, CH-8093 Zürich, Switzerland

^{§§} Department of Applied Physics, University of Athens, 36, Knossou Ano, Glyfada, 16561 Athens, Greece

LONG-TERM depletion of ozone has been observed since the early 1980s in the Antarctic polar vortex, and more recently at mid-latitudes in both hemispheres, with most of the ozone loss occurring in the lower stratosphere¹. Insufficient measurements of ozone exist, however, to determine decadal trends in ozone concentration in the Arctic winter. Several studies of ozone concentrations in the Arctic vortex have inferred that chemical ozone loss has occurred^{2–11}; but because natural variations in ozone concentration at any given location can be large, deducing long-term trends from

time series is fraught with difficulties. The approaches used previously have often been indirect, typically relying on relationships between ozone and long-lived tracers. Most recently Manney *et al.*¹¹ used such an approach, based on satellite measurements, to conclude that the observed ozone decrease of about 20% in the lower stratosphere in February and March 1993 was caused by chemical, rather than dynamical, processes. Here we report the results of a new approach to calculate chemical ozone destruction rates that allows us to compare ozone concentrations in specific air parcels at different times, thus avoiding the need to make assumptions about ozone/tracer ratios. For the Arctic vortex of the 1991–92 winter we find that, at 20 km altitude, chemical ozone loss occurred only between early January and mid February and that the loss is proportional to the exposure to sunlight. The timing and magnitude are broadly consistent with existing understanding of photochemical ozone-depletion processes.

During the 1991–92 winter, about 1,200 electrochemical cell (ECC) ozonesondes were launched as part of the European Arctic Stratospheric Ozone Experiment (EASOE) and from Canadian stations at middle and high northern latitudes, giving a uniquely detailed data set. Here we use the 10-day isentropic air-

parcel trajectories¹² calculated from the wind and temperature analyses of the European Centre for Medium-Range Weather Forecasts (ECMWF) to identify cases where ozone measurements were made at two different times in the same air mass. End-points for these trajectories included 26 sites. Figure 1 shows a map with the locations of the ozonesonde launching sites. Also shown is the trajectory of an air parcel which was probed by two ozonesondes at different times.

We have calculated the difference between the ozone amounts of each pair of measurements. For an individual match (see Fig. 1 legend for the definition of 'match' and for the criteria used), the uncertainty in the ozone and pressure measurements made by the ECC sondes¹³ and in the match condition itself is too large to allow us to draw firm conclusions. However, enough matches occurred during the 1991–92 winter yielding gaussian distributions to allow statistically valid conclusions to be drawn. Here we focus on 81 matches found between the 465 and 485 K isentropic surfaces (corresponding to a height of 20 km).

To a first approximation, air flows on isentropic surfaces. However, over the course of a few days some diabatic cooling occurs at high winter latitudes. The descent/ascent across isentropic surfaces is found for each trajectory from the cooling rates calculated in the UK Universities' Global Atmospheric Modelling Programme (UGAMP) global circulation model¹⁴, using the old radiative scheme¹⁵. An improved scheme has since been implemented, but the lower stratospheric cooling rates for clear skies in the sub-arctic winter compare well^{16,17}. The diabatic cooling along the trajectories was used to adjust the potential temperature of the ozone measurement at the end-point. In the absence of errors in the ozone measurements, the trajectory calculation or the cooling rate calculation, the difference in ozone mixing ratios at the two points in time is the chemical change in ozone along the trajectory.

Figure 2a shows the difference between pairs of ozone measurements as a function of the elapsed time. The results shown are for the period 4 January to 9 February 1992, and represent matches found in the edge and the core of the vortex. A substantial ozone decrease with time is observed with a loss rate of $0.044 \pm 0.007\%$ per flight hour (the error represents one standard deviation). Figure 2b shows how the ozone differences vary with the number of sunlit hours encountered. The calculated loss rate is $0.231 \pm 0.026\%$ per sunlit hour. Comparing the relative errors (ratios of the absolute errors to the central values) yields a 40% higher value for the loss rate per flight hour (0.16) than the loss rate per sunlit hour (0.11); this is reflected in the figures, as the increase of mixing ratio change is more uniform with time when plotting against sunlit rather than flight hours. Therefore the relation of ozone loss with the number of sunlit hours is stronger than that with the flight hours, and provides strong empirical evidence that chemical ozone loss requires light. This ozone loss rate is in reasonable agreement with those calculated using photochemical trajectory models for the same period^{18,19}.

Figure 3 shows how the ozone loss rate per sunlit hour varied during January and February 1992. Maximum loss rates were observed in middle and late January. Outside this period, no statistically significant ozone change was found, although it should be noted that in late December insufficient matches were found to draw any meaningful conclusions. Also shown in Fig. 3 is the mean vortex ozone depletion rate found by combining the sunlit ozone loss rate with the area of vortex exposed to sunlight. These empirically determined loss rates can be integrated to give an estimate of $38 \pm 4\%$ for the ozone loss which occurred in the edge and core of the vortex between the 465 and 485 K surfaces during the time. This integrated loss rate is larger than those found by empirical^{18,20} and model²¹ studies of measurements made during the Airborne Arctic Stratospheric Experiment AASE-II.

The errors stated so far are statistical in nature. Systematic errors could arise from the trajectory and the cooling-rate calcu-

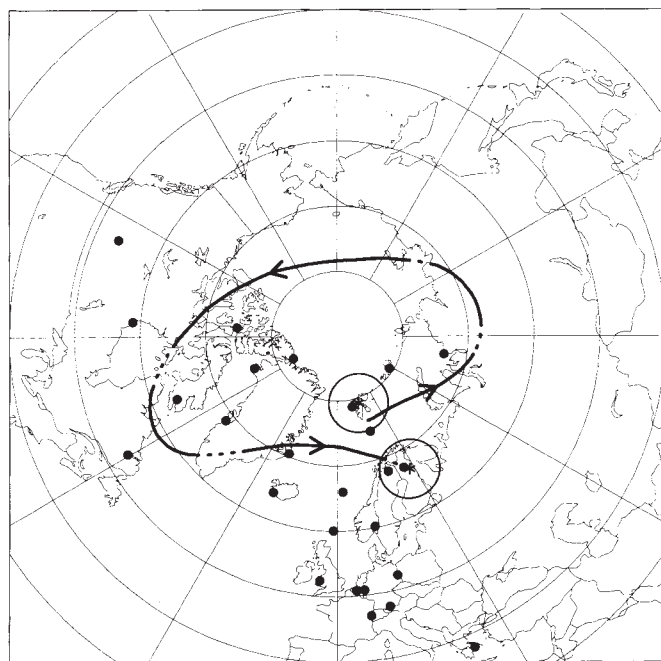


FIG. 1. Map of ozonesonde start sites (grey circles) and one example of a trajectory (475 K potential temperature surface) linking two ozonesonde sites and two soundings (24 January 1992, Ny-Ålesund in Spitsbergen; 29 January 1992, Sodankylä in Finland). Arrows show the direction in which the air parcel moved, and a dashed line indicates sunlit conditions. Potential temperature is the temperature an air parcel would reach if it were compressed adiabatically to 1000 hPa (1 atmosphere). A major source of uncertainty in the ozone loss analysis arises from the difference in time (and hence distance) between the ozone measurement and the trajectory passing over the station. The locations of the sondes when they reached the 475 K surface are marked by stars. These locations differ somewhat from the corresponding start sites because of wind drift. The start and end point of the trajectory (defined by the time the sondes reached the 475 K surface) are within the circles drawn with a radius of 500 km around the stars. Only cases where the sum of the distances between the ozone measurements and the trajectory at both ends is less than 500 km (like this case) were considered and are called matches. The 500-km limit was found by studying the standard deviation of the differences between the ozone mixing ratios at the start and the end site of the trajectories. A minimum plateau of the standard deviation was reached below 400 km. To increase the number of matches, the optimal compromise of 500 km was chosen.

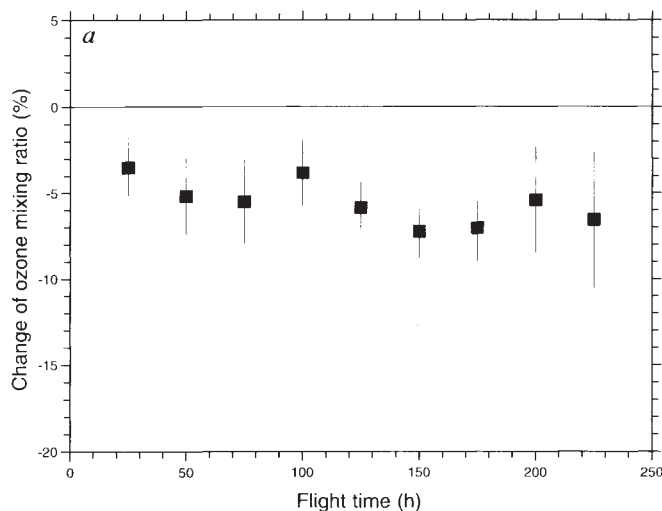
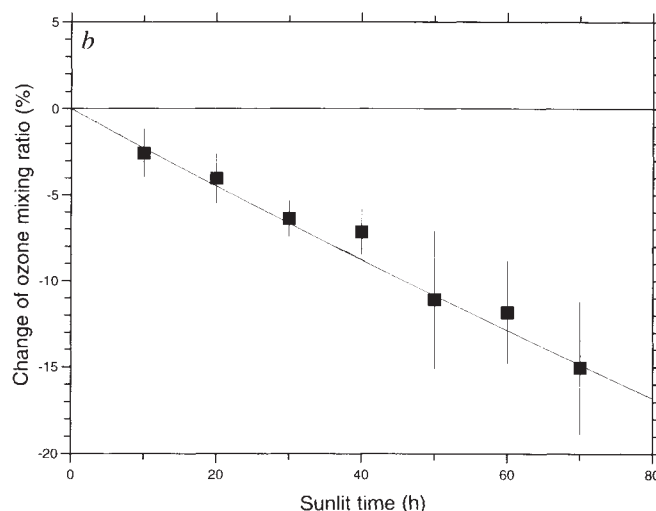


FIG. 2 *a*, Percentage change in ozone mixing ratio plotted against the number of hours between two measurements made in the same air mass at ~ 20 km altitude in the period 4 January to 9 February 1992. Between 6 and 24 matches contribute within ± 25 flight-time hours to each data point, with one exception of three matches at 200 flight-time hours. Measurements taken on the edge and in the core of the vortex are used. For the definition of the edge of the vortex, the quantity potential vorticity²² (PV) is used. Here the vortex edge is defined as $PV > 27 \text{ K m}^2 \text{ kg}^{-1} \text{ s}^{-1}$ on the 475 K isentropic surface. Error bars denote one standard deviation. *b*, Percentage change in ozone mixing ratio with the number of sunlit hours. Between 18 and 27 matches contribute within ± 10 sunlit hours to the first four data points, and between three and five matches contribute to the last three data points. An adequate exponential fit is shown by the solid line. A comparison of



this fit with *a* indicates that ozone loss takes place only during the sunlit hours. The observed loss with flight-time hours is only due to a correlation between the integrated sunlit time and the flight time. Note that the largest total loss of ozone is found for matches which received the most sunlight (right-hand side). These matches are redistributed over a wide range of flight-time hours on the right of *a*. By taking the means, only a few of them combine with matches that got only a few sunlit hours during a long flight time. As a result, this ensemble of matches experienced less sunlit hours than the other matches in proportion to their flight times. Therefore the values on the right-hand side of *a* are much less negative. Owing to a statistical fluctuation which is not significant, the matches on the left of *a* with shorter flight times show more ozone change than those with longer flight times in proportion to their flight times. In *b*, both types combine in the left-hand data points.

lations, but are unlikely to be present in the measurements. At the 475 K level, the ozone mixing ratio is generally higher inside the vortex than outside. A systematic error in the calculation of the trajectories which caused a false shift across the vortex edge could thus lead to an incorrect ozone loss rate. There are two arguments which suggest that this did not affect our study. First, there is no reason to suppose that such processes would have a better relation with sunlit hours than with time of flight. Second, the potential vorticity²² (PV) can be used as a dynamical tracer for this purpose because it is known to be conservative (stable) on a timescale of 1–2 weeks and also because it decreases from

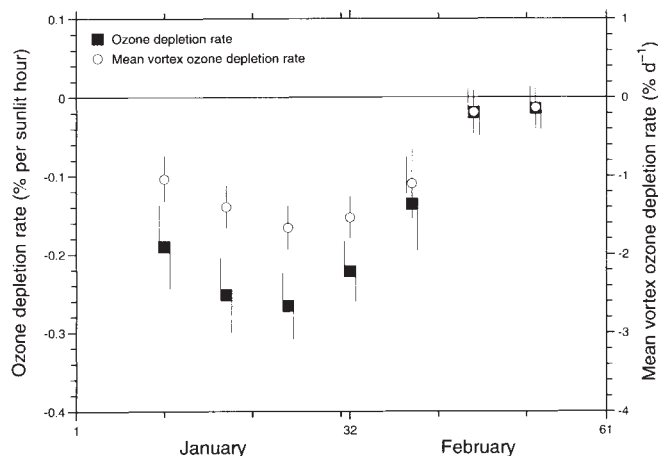


FIG. 3 Time evolution of the ozone depletion rate (% per sunlit hour) and of the mean vortex ozone depletion rate (% per day) during January and February 1992. Between 13 and 30 matches contribute within ± 7 days to each data point. Otherwise, conditions are as Fig. 2*a*.

the inside to the outside of vortex regions. The mean difference between PV at the start and the end sites of the trajectories relative to the mean PV is small ($+0.8\% \pm 1.5\%$), so a systematic drift of the trajectories to either side is unlikely to be important.

The ozone mixing ratio increases with altitude in the lower stratosphere, and so any overestimate of the diabatic cooling rate would also lead to an overestimate in the ozone losses. The analyses have been repeated with 150%, 50% and 0% of the model cooling rates in the time period given in Fig. 2*b*. Whereas the 150% run yields a loss rate of $0.27 \pm 0.03\%$ per sunlit hour, the 50% run still yields a loss rate of $0.16 \pm 0.03\%$ per sunlit hour. Even if no cooling is assumed (0%), the loss rate is about half of the 100% case ($0.12 \pm 0.05\%$ per sunlit hour). Also in the last case we notice an apparent ozone production of $0.07 \pm 0.03\%$ per sunlit hour in the middle of February. There is no chemical explanation for such an effect, which should really be taken as evidence that cooling and descent of air did take place. Furthermore our cooling rates, which average $\sim 0.7 \text{ K d}^{-1}$, are in good agreement with recent reported other model descent rates²³ ($\sim 0.75 \text{ K d}^{-1}$) and N_2O tracer observations²⁴ (0.7 K d^{-1}).

The meteorology of the 1991–92 winter has been described elsewhere^{25–27}. Briefly, temperatures cold enough to form polar stratospheric clouds (PSCs) existed intermittently, but with increasing frequency, from early December until mid January. A minor warming then occurred, which inhibited further large-scale formation of PSCs, but it is possible that some PSCs were formed locally, as visual sightings were reported over the Scandinavian mountains in conditions suitable for the formation of orographic waves²⁸. CIO measurements made by the Upper Atmosphere Research Satellite Microwave Limb Sounder instrument showed high values of CIO (up to 2.5 p.p.b.v. at 19 km altitude) in the vortex²⁹ until the satellite turned to look south on 13 January 1992. When the instrument next looked north in mid February, CIO levels had halved. Other measurements made

during EASOE and from the NASA ER-2 research aircraft similarly show that the main period of chlorine activation coincided with, and slightly lagged the cold temperatures^{30–32}. The time evolution of the ozone loss shown in Fig. 3 is thus consistent both with our understanding of chlorine- (and bromine-) induced ozone loss obtained empirically from observations, and with the results of calculations using photochemical models. □

Received 12 September 1994; accepted 29 March 1995.

- Harris, N. R. P. et al. *Scientific Assessment of Ozone Depletion: 1994* Ch. 1 (WMO Global Ozone Research and Monitoring Project, Report No. 37) (World Meteorological Organisation, Geneva, 1995).
- Hofmann, D. J. et al. *Nature* **340**, 117–121 (1989).
- Schoeberl, M. R. et al. *Geophys. Res. Lett.* **17**, 469–472 (1990).
- McKenna, D. S. et al. *Geophys. Res. Lett.* **17**, 553–556 (1990).
- Hofmann, D. J. & Deshler, T. *Nature* **349**, 300–305 (1991).
- Koike, M. et al. *Geophys. Res. Lett.* **18**, 791–794 (1991).
- Kyrö, E. et al. *J. Geophys. Res.* **97**, 8083–8091 (1992).
- Proffitt, M. H. et al. *Science* **261**, 1150–1154 (1993).
- Braathen, G. O. et al. *Geophys. Res. Lett.* **21**, 1407–1410 (1994).
- Larsen, N., Knudsen, B. M., Mikkelsen, I. St., Jørgensen, T. S. & Eriksen, P. *Geophys. Res. Lett.* **21**, 1611–1614 (1994).
- Manney, G. L. et al. *Nature* **370**, 429–434 (1994).
- Knudsen, B. M. & Carver, G. D. *Geophys. Res. Lett.* **21**, 1199–1202 (1994).
- Kerr, J. B. et al. *Atmosphere-Ocean* **32**, 685–716 (1994).
- Carver, G. D., Norton, W. A. & Pyle, J. A. *Geophys. Res. Lett.* **21**, 1451–1454 (1994).

- Geleyn, J. F. & Hollingsworth, A. *Beitr. Phys. Atmos.* **52**, 1–16 (1979).
- Morcrette, J. J. *Tech. Memo. 165* (Res. Dep. Eur. Cent. for Medium Range Weather Forecasts, Reading, UK, 1989).
- Morcrette, J. J. *J. Geophys. Res.* **96**, 9121–9132 (1991).
- Lutman, E. R., Toumi, R., Jones, R. L., Lary, D. J. & Pyle, J. A. *Geophys. Res. Lett.* **21**, 1415–1418 (1994).
- Müller, R. et al. *Geophys. Res. Lett.* **21**, 1427–1430 (1994).
- Browell, E. V. et al. *Science* **261**, 1155–1158 (1993).
- Salawitch, R. J. et al. *Science* **261**, 1146–1149 (1993).
- Hoskins, B. J., McIntyre, M. E. & Robinson, A. W. Q. *J. R. met. Soc.* **111**, 877–946 (1985).
- Rosenfield, J. E., Newman, P. A. & Schoeberl, M. R. *J. Geophys. Res.* **99**, 16677–16689 (1994).
- Strahan, S. E., Rosenfield, J. E., Loewenstein, M., Podolske, J. R. & Weaver, A. J. *Geophys. Res.* **99**, 20713–20723 (1994).
- Farman, J. C., O'Neill, A. & Swinbank, R. *Geophys. Res. Lett.* **21**, 1195–1198 (1994).
- Naujokat, B., Petzoldt, K., Labitzke, K. (Met. Inst. report Beilage zur Berliner Wetterkarte, SO 18/92, Berlin, 1992).
- Newman, P. A. et al. *Science* **261**, 1130–1158 (1993).
- Godin, S. et al. *Geophys. Res. Lett.* **21**, 1335–1338 (1994).
- Waters, J. W. et al. *Nature* **362**, 597–602 (1993).
- Geophys. Res. Lett.* **21**, 1189–1490 (1994).
- Geophys. Res. Lett.* **20**, 2499–2578 (1993).
- Science* **261**, 1130–1158 (1993).

ACKNOWLEDGEMENTS. We thank the following persons and institutions for providing their data and useful comments: H. Claude, Met. Obs. Hohenpeißenberg, Germany; U. Görsdorf, Met. Obs. Lindenberg, Germany; P. A. Hughes, UK Met. Office, Bracknell, UK; S. J. Reid, Univ. of Wales, Aberystwyth; V. V. Rudakov and I. Zaitcev, CAO, Moscow, Russia; H. G. Smit, KFA-Jülich, Germany; and the Swedish Space Corporation at Esrange, Kiruna. This work was supported by the Bundesministerium für Forschung und Technologie, DG XII of the EC, the UK Department of Environment, the UK Natural and Environmental Research Council and the Danish Research Academy.

Archaean cratonic roots, mantle shear zones and deep electrical anisotropy

M. Mareschal*, R. L. Kellett*†, R. D. Kurtz‡, J. N. Ludden§, S. Ji|| & R. C. Bailey¶

* Gémie Minéral, École Polytechnique, CP 6079, succ. Centre-ville, Montréal, Québec H3C 3A7, Canada

† Geology Department, University of New Brunswick, PO Box 4400, Fredericton, New Brunswick E3B 5A3, Canada

‡ Geological Survey of Canada, 1 Observatory Crescent, Ottawa, Ontario K1A 0Y3, Canada

§ CRPG-CNRS, Vandoeuvre-les-Nancy, 54501 France

|| Département de Géologie l'Université de Montréal, CP 6128, succ. Centre-ville, Montréal, Québec H3C 3J7, Canada

¶ Physics Department, University of Toronto, Toronto, Ontario M5S 1A7, Canada

THE extent to which the mantle participated in the growth and stabilization of ancient cratons is central to our understanding of the evolution of the continents¹. The detection of seismic anisotropy beneath Precambrian North America, for example, has been interpreted as showing that strain-induced orientation of mantle minerals in subcontinental lithospheric mantle can preserve a record of ancient episodes of deformation². Here we present magnetotelluric measurements from the Superior Province of the Canadian shield, which reveal pronounced electrical anisotropy in the upper 100 km of the underlying mantle. We argue that this anisotropy is best explained by conducting graphite films, oriented within fractures or on grain boundaries, and associated with metasomatism of the mantle roots of major Archaean shear zones which transect the entire Superior Province. The uppermost mantle beneath the Canadian shield has therefore remained fixed to the crust and isolated from significant tectonic reworking since the late Archaean.

Magnetotellurics, a deep-sounding electromagnetic technique based on measurements of natural electric and magnetic fields at the surface of the Earth, is well suited to probing the electrical response of the lithosphere. Because the upper crust of the Canadian shield is very resistive³ (>10,000 Ω m), the conven-

tional period range (0.02–2,000 s) resolves electrical features from a few kilometres to at least 100 kilometres depth. Recent developments in instrumentation and data processing yielding accurate evaluation of the phase between the electric and magnetic fields, and the application of new impedance analysis methods^{4,5}, have significantly reduced the uncertainties associated with previous electrical models of the Earth's structure⁶. Data processing, modelling and inversion of two subsets of the data used in this study have been published elsewhere^{7,8} and are representative of the processing applied to all of the data.

Although the magnetotelluric soundings span an area of approximately 500 × 400 km, and cross several geological sub-provinces of the Superior Province (Fig. 1, west of 78° W), the data are remarkably uniform in character. Most soundings can be characterized by a simple resistivity model consisting of a series of one-dimensional crustal layers, overlying a two-dimensional layer, identified by the separation of the phase curves for two orthogonal directions, at periods between 10 and 500 s (Fig. 2). Specific, vertical-magnetic-field analyses^{9,10} suggest that the phase split cannot be due to induction in discrete bodies, leaving anisotropy as the most realistic cause of the magnetotelluric response. However, the most intriguing feature of Fig. 1 is the manner in which the azimuth of enhanced conductivity varies over horizontal distances of the order of a few hundred kilometres, from approximately N80E in the Pontiac, eastern Abitibi and northern Grenville, to approximately N20W in northern Kapuskasing and eastern Wawa.

Inversion of the magnetotelluric data indicates that the crustal structure is typical of Precambrian shields⁶, with a low-resistivity layer (layer 3, Fig. 2b, d) at lower-crustal depths. The top of the anisotropic layer (layer 4 in Fig. 2b, d) is near 45 km in the Pontiac, eastern Abitibi and northern Grenville (Fig. 1), but closer to 65 km under the Kapuskasing uplift, a change in depth that parallels the increase in Moho depth determined from seismic data¹¹. The base of the anisotropic layer is not well resolved by these magnetotelluric data but cannot be deeper¹² than 150 km and the ratios of orthogonal horizontal resistivities can be as high as 1:15.

In contrast to seismic anisotropy, the low resistivity and large electrical anisotropy in the upper mantle cannot be explained by lattice preferred orientation of typical mantle minerals¹³. Rather, electrical conduction is achieved through minor constituents such as graphite or saline fluids, and the anisotropy must arise

from the geometry of the interconnection between these conductive phases. This reflects either present-day mantle flow, mantle flow at the time of major crustal deformation (late Archaean) or a more recent deformation of the lower lithosphere that did not affect the crust.

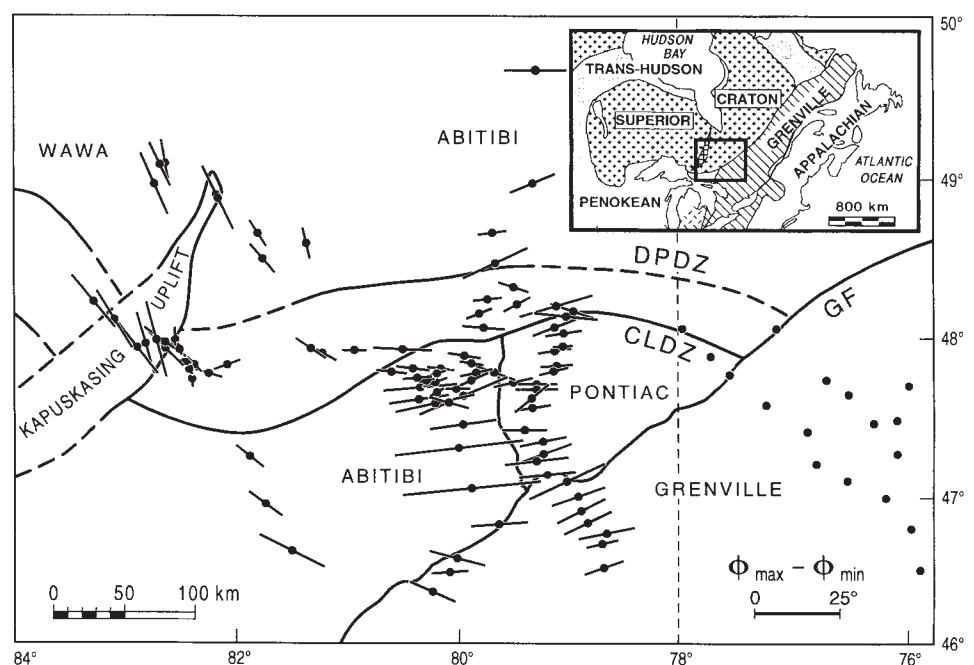
Seismic tomography of the mantle suggests that anomalously cold, deep (up to 250 km) roots exist under the Canadian shield¹⁴, and rheological profiles (Fig. 3) confirm that the mantle is at present brittle down to a depth of ~150 km, in good agreement with the depth extent of the electrical anisotropy. Thus there is no mantle flow at present in the range 50–150 km depth, and the anisotropy cannot be explained by the infiltration of saline fluids resulting from deeper, modern convection systems. Further, one would expect that a recent source of anisotropy should be at least partially reflected in the present stress field. But the current pattern of maximum horizontal stress over the Canadian shield is uniform¹⁵ (N75E) and bears no relation to the variations in azimuth of the electrical anisotropy (Fig. 1, west of 78° W). Similarly, the pattern of Proterozoic dykes¹⁶, potentially mapping a palaeostress field at that time, does not correlate with the azimuths shown in Fig. 1. Late Proterozoic and Palaeozoic rifting has occurred in the Canadian shield, but the two branches closest to the survey area (Ottawa–Bonnechere graben, and the Lake Timiskaming fault system) are shallow structures with a diffuse pattern of seismicity¹⁷ which, again, produce no perturbation in the regional stress field, and thus will not have a significant effect on magnetotelluric data at periods of 100 s.

If the anisotropy is of large scale (as indicated by the results of Fig. 1), it seems implausible that the upper mantle could develop a more recent geological structure without significant effects on the overlying Archaean crust. An explanation in terms of large-scale mantle structure of Archaean age is therefore a real possibility. The geothermal conditions during the Archaean have been subject to considerable discussion; however, it appears that upper-mantle temperatures may have been only 50–100 °C higher than at present^{18,19}, with Moho temperatures in the south-western Superior Province¹⁸ of the order of 800 °C at 2,500 Myr ago. Such temperatures would have allowed both the generation and the preservation of electrical anisotropy if the conductive phase were graphite²⁰, but render it impossible if it were water (dewatering would have occurred in less than 10⁵ years, ref. 16). Further, observations on granulites indicating an anomalously

low degree of melting²¹ and a high concentration of grain-boundary graphite²², imply a rather low concentration of Archaean hydrous fluids at lower-crustal levels. Assuming that most of these fluids were produced by degassing of the upper mantle, their anomalous concentration tends to support a carbon-related source for the upper-mantle electrical anisotropy, the minute quantities of graphite necessary to explain the present-day conductivity having easily been deposited in organized cracks or grain boundaries. Note that during the most active phases of metamorphism (~2,680 Myr), the passage of fluids would render rocks locally brittle even at temperatures higher than 900 °C, with hydrofracturing along foliation planes relieving pressure and consequently helping the precipitation of the conductive phase. The temperature decrease until 2,500 Myr ago would have helped preserve the anisotropic signature, even as activity along the faults started to decrease, the presence of a thick but light refractory lithosphere providing protection from higher mantle temperatures¹⁹.

Whereas the western Superior Province is characterized by long parallel belts striking east–west, the Abitibi subprovince is a collage of small irregular-shaped terranes²⁴. Correlating the azimuth of electrical anisotropy with individual greenstone-belt boundaries proved to be unsuccessful (Fig. 1), but the pattern of conductive directions bears resemblance to the traces of two major Archaean fault zones. The Cadillac–Larder Lake (CLDZ) and Destor–Porcupine (DPDZ) deformation zones (Fig. 1) are continuous across southern Abitibi; they have been correlated with similar faults in the Wawa Subprovince, and with shear zones of granulite facies in the Kapuskasing uplift²⁵, where their traces have been offset by Proterozoic faulting. There, regional ages in the range 2,680–2,500 Myr overlap those associated with a period of gold deposition in the lower-grade fault zones of Abitibi²⁶, another piece of indirect evidence that carbonate-rich fluids may well represent small, deeper partial melts of mantle which, as temperatures decreased, were trapped in the lithospheric roots. The fact that no geophysical investigation has imaged fossil connection between the upper-crustal fault zones and shear zones in the mantle does not preclude such a relation. The lower crust certainly stayed soft long after the Archaean: the rheological models on Fig. 3, the uniform conductor (layer 3) in Fig. 2, and the seismic reflection sections across the DPDZ and the CCLZ (in which the faults appear to become listric at

FIG. 1 Variations in azimuth and amplitude of the upper-mantle electrical anisotropy at the sounding sites (filled circles) over the study area (81° to 78° W). The azimuth of each solid bar indicates the most conductive direction, and the length of the bar is proportional to the phase difference between this direction and the orthogonal direction (see Fig. 2). These features are schematically plotted in two different sets that are close to the periods at which the maximum split occurs in the phases, (42 s east of 81° W, and 113 s west of 81° W). Filled circles east of 78° W represent soundings for which the anisotropic response is either highly disturbed (four Abitibi sites near the high-angle intersection of the DPDZ, CLDZ and GF) or absent (Grenville Province). The major tectonic features of the southern Superior Province are shown as simple lineaments (DPDZ, Destor–Porcupine deformation zone; CLDZ, Cadillac–Larder Lake deformation zone; GF, Grenville front). Inset, location of the survey area within the Canadian shield, showing the Archaean cratons and the surrounding Proterozoic and Palaeozoic fold belts.



intermediate-crustal level²⁷) are all consistent with a long-lasting ductile lower crust. However, this does not imply that upper-crustal and upper-mantle tectonics are disconnected, as recently shown for the active Tibet²⁸. Moreover, seismic reflections recorded just north of our study area²⁹, although discontinuous across a lower-crustal zone of horizontal reflectivity, reveal a relict Archaean suture extending from the upper crust to at least 30 km into the mantle, suggesting that even during the Archaean, horizontal transmission of stress from a common lateral source may have been more important than vertical stress coupling. Thus we contend that the fault zones shown in Fig. 1 extended from the surface to at least lower-crustal depth during the late Archaean, and that they probably provided the conduits necessary for migration of CO₂-rich fluids from the upper mantle.

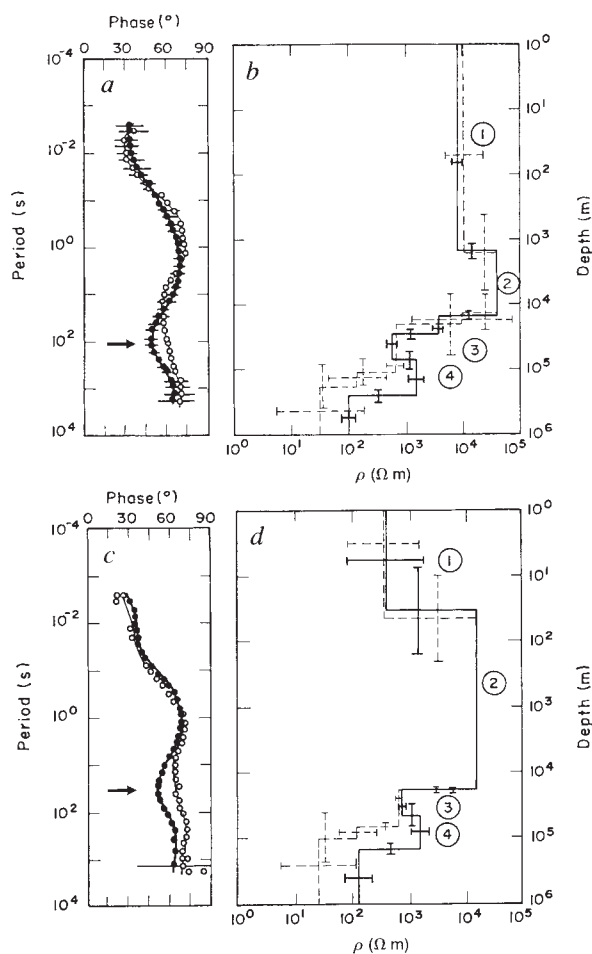


FIG. 2 Examples of magnetotelluric sounding curves (left) and resistivity models (right) for the Kapuskasing uplift (a, b) and the Pontiac Subprovince (c, d). The phase curves are shown in the most conductive direction (higher phases) at upper-mantle level (open circles; N25W in a and N80E in c) and orthogonal directions (filled circles). The arrows indicate the periods of maximum anisotropy used in Fig. 1. The data are remarkably uniform in character and can be interpreted in terms of a one-dimensional resistivity structure at periods <10 s. At all sites, a reduction in resistivity is observed at lower-crustal level (layer 3 in b and d) which is typical of Precambrian terranes⁶. The anisotropic layer (4) is found in the upper mantle (dashed line corresponds to the most conductive direction; solid line corresponds to the resistivity in the orthogonal direction). The error bars on the resistivity and depth of each layer represent one standard deviation misfit on the magnetotelluric sounding curves, and the model response is shown as solid lines in a and c. At some more disturbed sites in the Kapuskasing uplift, static distortion of the magnetic as well as the electric fields by isolated shallow bodies can account for part of the phase split¹⁰, but this cannot explain the spatial uniformity of the magnetotelluric responses in Fig. 1.

It could be argued that the electrical anisotropy simply reflects the preferred grain boundary orientation in the entire lithospheric mantle beneath the Superior Province; during plastic deformation of the upper mantle, graphite would deform into continuous thin films parallel to lineation and foliation. This cannot be resolved from the magnetotelluric observations. However, the spatial uniformity of the magnetotelluric responses implies that, if the anisotropy is due to graphite concentration in the mantle shear zones, these need to be of the order of 50 km wide at upper-mantle levels, with a maximum separation of ~100 km. This pattern is consistent both with geological models³⁰ and with the present distribution of major upper-crustal faults³¹ in the southeastern Superior Province. Moreover, our interpretation establishes a connection between the more restrictive presence of conductive material in the deep shear zones and the alkaline rocks and gold deposits found in shallower parts of the deformation zones²⁶.

In addition to the relation between conductive azimuths and crustal shear zones, the magnetotelluric responses also help constrain the lithospheric structure of the Canadian shield. The anisotropy remains constant across the Grenville front (GF), south of the Pontiac, which we interpret as being caused by the Archaean upper-mantle fabric extending beneath the Grenville parautochthonous terranes. This model is consistent with the persistence of Pontiac-style lower-crustal seismic reflectivity and Archaean peak-metamorphic ages, for up to 35 km into this part

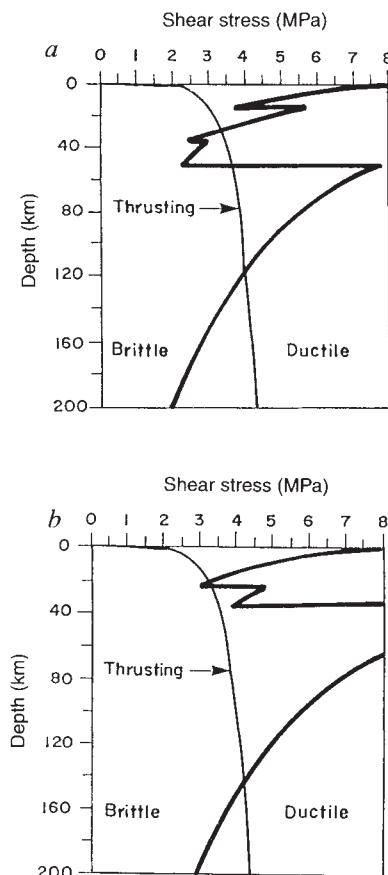


FIG. 3 Rheological models of present crustal and upper-mantle sections (the curve labelled 'thrusting' indicates the present state of stress in the study area) for the Kapuskasing uplift (a) and the Pontiac Subprovince (b). The thick black line represents the boundary between brittle and ductile behaviour. The models are based on local lithological stratifications and seismic reflection profiles, temperature and heat flow data, and friction and creep laws; they suggest that the mantle is at present brittle down to ~150 km.

of the Grenville Province³². Note also that if the temperatures were not reset by the Grenville orogeny, the local geometry of the faults (the GF, CLDZ and DPDZ all being parallel) would have helped enhance anisotropy. Further east, the CLDZ and the DPDZ are truncated (at a high angle) by the GF, resulting in attenuated upper-mantle anisotropy at adjacent sites in Abitibi. The terranes in the Grenville Province south of the GF, although originally Archaean, are strongly affected by late Proterozoic metamorphism³², and the lack of electrical anisotropy indicates that the mantle was probably reset (although higher crustal conductivities may be screening the upper-mantle

responses). In contrast, in the western part of the study area, the anisotropy remains strong but changes direction, suggesting that the upper-mantle fabric may have been significantly affected, but not destroyed, by the intra-cratonic compression that led to Kapuskasing uplift.

Regardless of its precise origin, the survival of electrical anisotropy in the upper mantle of the Canadian shield adds to the growing body of evidence that the Archaean cratons developed a upper-mantle root relatively early in their accretion history and that in the case of the Superior Province, this root has survived relatively intact to modern times. □

Received 26 August 1994; accepted 29 March 1995.

- Hoffman, P. F. *Phil. Trans. R. Soc. A* **331**, 523–532 (1990).
- Silver, P. G. & Kaneshima, S. *Geophys. Res. Lett.* **20**, 1127–1130 (1993).
- Bailey, R. C., Craven, J. A., Macnae, J. C. & Polzer, B. D. *Nature* **340**, 136–138 (1989).
- Groom, R. W. & Bailey, R. C. *J. geophys. Res.* **94**, 1913–1925 (1989).
- Bahr, K. *Phys. Earth planet. Inter.* **66**, 24–38 (1991).
- Jones, A. G. in *Continental Lower Crust* (eds Fountain, D., Arculus, R. & Kay, R.) 81–143 (Elsevier, Amsterdam, 1992).
- Kellett, R. L., Mareschal, M. & Kurtz, R. D. *Geophys. J. Int.* **111**, 141–150 (1992).
- Kurtz, R. D., Craven, J. A., Niblett, E. R. & Stevens, R. A. *Geophys. J. Int.* **113**, 483–498 (1993).
- Zhang, P., Pedersen, L. B., Mareschal, M. & Chouteau, M. *Geophys. J. Int.* **113**, 693–700 (1993).
- Chave, A. D. & Smith, J. T. *J. geophys. Res.* **99**, 4669–4682 (1994).
- Clowes, R. M. et al. *Can. J. Earth Sci.* **29**, 1813–1864 (1992).
- Schultz, A., Kurtz, R. D., Chave, A. D. & Jones, A. G. *Geophys. Res. Lett.* **20**, 2941–2944 (1993).
- Shankland, T. & Duba, A. *Geophys. J. Int.* **103**, 25–31 (1990).
- Jordan, T. H. *Rev. Geophys. Space Phys.* **13**, 1–12 (1975).

- Zoback, M. L. *J. geophys. Res.* **97**, 11761–11782 (1992).
- West, G. F. & Ernst, R. E. *Can. J. Earth Sci.* **28**, 1797–1811 (1991).
- Forsyth, D. A. *Can. J. Earth Sci.* **18**, 103–119 (1981).
- Percival, J. in *Archean Crustal Evolution* (ed. Condie, K. C.) 357–410 (Elsevier, Amsterdam, 1994).
- Nisbett, E. G., Cheadle, M. J., Arndt, N. T. & Bickle, M. T. *Lithos* **30**, 291–307 (1993).
- Meissner, R. & Strehlau, T. *Tectonics* **1**, 73–89 (1982).
- Bailey, R. C. *Geophys. Res. Lett.* **17**, 1129–1132 (1990).
- Skippen, G. B. & Marshall, D. D. *Can. Mineralogist* **29**, 693–705 (1991).
- Mareschal, M., Fyfe, W. S., Percival, J. & Chan, T. *Nature* **357**, 674–676 (1992).
- Jackson, S. & Fyon, T. *Geology of Ontario* (OGS Vol. 4, Part 1) (Ontario Geological Survey, Sudbury, 1991).
- Leclair, A., Ernst, R. E. & Hattori, K. *Geology* **21**, 399–402 (1993).
- Feng, R., Kerrich, R., McBride, S. & Farrar, E. *Can. J. Earth Sci.* **29**, 1389–1411 (1992).
- Green, A. G. et al. *Nature* **344**, 327–330 (1990).
- Jin, Y., McNutt, M. & Zhu, Y. *Nature* **371**, 669–674 (1994).
- Calvert, A. J., Sawyer, E., Davis, W. & Ludden, J. *Nature* (submitted).
- Nicolas, A. *Phys. Earth planet. Inter.* **78**, 337–342 (1993).
- Lacroix, S. & Sawyer, E. *Can. J. Earth Sci.* (in the press).
- Kellett, R. L., Barnes, A. E. & Rive, M. *Can. J. Earth Sci.* **31**, 282–292 (1994).
- Indares, A. & Martignole, J. *Can. J. Earth Sci.* **27**, 371–386 (1990).

Absorption of ant-provided carbon dioxide and nitrogen by a tropical epiphyte

Kathleen K. Treseder*, Diane W. Davidson & James R. Ehleringer

Department of Biology, University of Utah, Salt Lake City, Utah 84112, USA

ALTHOUGH ant-plant mutualisms have been described in many ecosystems, the magnitude of the direct benefits from such relationships are hard to quantify. In Bako National Park, Sarawak, Malaysia, stunted 'kerangas' forests occur on nutrient-poor sandstone hills^{1–3}. As trees are widely spaced and have a sparse leaf area, a significant amount of light reaches the tree trunks and enables a diverse community of epiphytes to thrive there⁴. One of these epiphytes, *Dischidia major* (Vahl) Merr. (Asclepiadaceae), has evolved unusual methods for enhancing carbon and nitrogen acquisition. We show here that a mutualistic relationship exists between ants of the genus *Philidris* and their host, *D. major*. Using stable isotope analysis, we calculate that 39% of the carbon in occupied host plant leaves is derived from ant-related respiration, and that 29% of the host nitrogen is derived from debris deposited into the leaf cavities by ants.

In addition to small coin-shaped leaves, *D. major* has evolved sac-like 'ant leaves' (Fig. 1), in which ants of the genus *Philidris*⁵ (Dolichoderinae) frequently raise young and deposit debris (faeces, dead ants and scavenged insect parts)^{4,6}. Adventitious roots from *D. major* grow through the cavity opening and proliferate wherever debris has accumulated⁴. It has been proposed

that the *Dischidia-Philidris* relationship is mutualistic and that *D. major* uses ant debris as a nitrogen source⁴; also, the stomata on the internal surfaces of leaf cavities^{7–9} may absorb ant-respired carbon dioxide and thereby reduce transpirational water loss^{8,9}, but neither of these suggestions is supported by experimental evidence.

We tested these hypotheses by measuring stable isotope ratios ($\delta^{13}\text{C}$, $\delta^{15}\text{N}$) of ants, hosts and substrates, and by capitalizing on differences in the isotope composition of possible nutrient sources. This approach enabled us to quantify the benefit to hosts from their symbiotic ants.

The $\delta^{13}\text{C}$ value of *Dischidia* depends on both the $\delta^{13}\text{C}$ of the source CO_2 ($\delta^{13}\text{C}_{\text{CO}_2}$) and discrimination during carbon fixation (Δ), thus $\delta^{13}\text{C}_{\text{plant}} = \delta^{13}\text{C}_{\text{CO}_2} - \Delta$. *D. major* has obligate crassulacean acid metabolism (CAM)^{10,11}, and we found that uninhab-

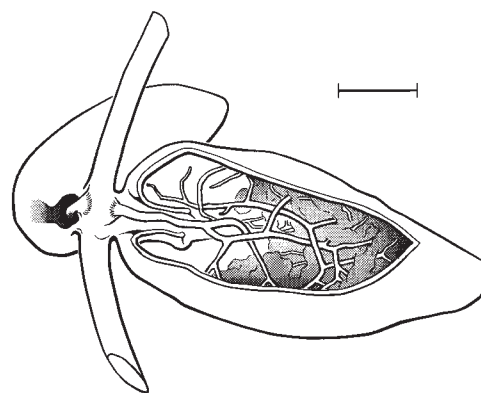


FIG. 1 Two ant-occupied leaves of *Dischidia major*: one is cut away to show adventitious roots and ant debris. The modified ant leaves are rolled up to form an enclosed abaxial surface and a cavity accessible through a small basal opening near the petiole^{4,6}. Scale bar, 1 cm.

* Present address: Department of Biological Sciences, Stanford University, Stanford, California 94305, USA.

ited plants had a mean $\delta^{13}\text{C}_{\text{plant}}$ value of -16.0‰ , typical of CAM plants¹². Given an atmospheric $\delta^{13}\text{C}_{\text{CO}_2}$ value of -7.9‰ (ref. 13), Δ is calculated as 8.1‰ .

Ant respiration and decomposing debris are additional possible sources of CO_2 in ant-occupied leaves and could alter $\delta^{13}\text{C}_{\text{CO}_2}$. The carbon isotope ratio of animals (and their respiration) is essentially the same as that of their food^{14,15}. *Philidris* ants feed on the exudates of Homoptera⁴ that ingest the phloem of C_3 rainforest trees. Subsequently, the ants have low $\delta^{13}\text{C}$ values averaging -25.9‰ (Fig. 2). $\delta^{13}\text{C}_{\text{CO}_2}$ used for photosynthesis by *D. major* could vary from -7.9‰ , if no ant-derived CO_2 is taken up, to -25.9‰ , if ant-derived CO_2 accounts for all fixed CO_2 . The corresponding $\delta^{13}\text{C}_{\text{plant}}$ values would be expected to vary from -16‰ to -34‰ .

If *D. major* incorporates ^{13}C -deficient carbon dioxide respired by the ants or their debris, colonized leaves should have lower $\delta^{13}\text{C}$ values than do uncolonized leaves. In addition, as colonies deposit debris inside a leaf only after they have raised brood there for some time⁴, ant leaves with debris should have had a longer time to accumulate ant-respired CO_2 . Overall, completely vacant leaves should have the highest $\delta^{13}\text{C}$ values (about -16.0‰), followed by ant-occupied leaves that are debris-free, and finally, by debris-filled leaves. The three groups followed the expected pattern (Fig. 2). Therefore individuals of *D. major* take up significant amounts of carbon from *Philidris*.

To estimate the fraction of plant carbon derived from ant-related respiration (per cent CO_2 from ants), we used the two-member mixing model:

$$\% \text{CO}_2 \text{ from ants} = \frac{\delta^{13}\text{C}_{\text{plant}} + \Delta - \delta^{13}\text{C}_{\text{CO}_2, \text{atm}}}{\delta^{13}\text{C}_{\text{ant}} - \delta^{13}\text{C}_{\text{CO}_2, \text{atm}}} \times 100\%$$

where $\delta^{13}\text{C}_{\text{CO}_2, \text{atm}}$ and $\delta^{13}\text{C}_{\text{ant}}$ are the isotope compositions of the atmosphere and ants, respectively. We calculated the fraction of plant carbon derived from ant-related respiration as 39% ($\pm 8\%$, $n=5$) for ant-occupied, debris-filled leaves; 27% ($\pm 4\%$,

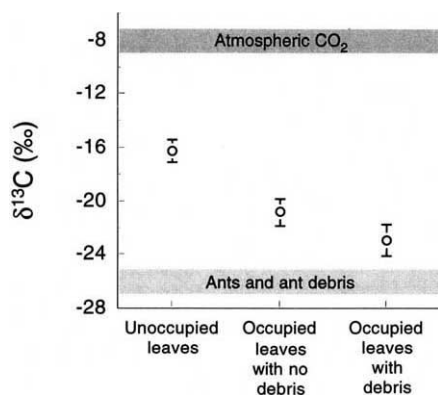


FIG. 2 Mean carbon isotope values ± 1 s.e. for *Dischidia major* ant leaves and the potential CO_2 sources for these leaves (atmospheric CO_2 (-7.9‰), ants ($-25.9\text{‰} \pm 0.3$, $n=6$), or ant debris ($-25.9\text{‰} \pm 0.9$, $n=3$); shaded bars). Values differed significantly for the three groups of ant leaves ($H_{[5,6]}=9.510$, $P<0.009$). $\delta^{13}\text{C}$ values of ant-related debris, a possible source of carbon dioxide, did not differ significantly from those of the ants ($U_{[3,6]}=12$, $P=0.731$).

METHODS. Material was collected along the Lintang trail in Bako National Park, Sarawak, Malaysia in August 1993. With an isotope ratio mass spectrometer (Delta S, Finnigan MAT at SIRFER, University of Utah, Salt Lake City), isotope ratios of ants, leaves and ant debris were measured on 2–3 mg samples. Isotope compositions are reported in $\delta^{13}\text{C}$ values as $\delta^{13}\text{C} = [(^{13}\text{C}/^{12}\text{C}_{\text{sample}})/(^{13}\text{C}/^{12}\text{C}_{\text{standard}}) - 1] \times 1,000\text{‰}$, where $^{13}\text{C}/^{12}\text{C}_{\text{sample}}$ and $^{13}\text{C}/^{12}\text{C}_{\text{standard}}$ are the isotope ratios of the sample and standard, respectively²⁵.

$n=6$) for ant-occupied leaves without debris; and 4% ($\pm 4\%$, $n=6$) for vacant leaves.

For ant-related CO_2 to contribute so significantly to the carbon balance of the plant, the CO_2 concentration inside ant-occupied leaves should be raised above atmospheric values. By having interior stomata fed by an increased CO_2 concentration, the plant probably significantly increases its water-use efficiency (photosynthetic carbon gain to transpirational water loss). The capture of ant-respired CO_2 may increase carbon gain (enhancing growth) and also reduce transpiration by curbing stomatal activity. Moreover, the interior stomata transpire into a partially enclosed cavity, where relative humidity should be higher than the surrounding atmosphere, resulting in further reduction in transpiration. Although *D. major* grows in a tropical climate, water loss is a concern because the epiphytes have no access to soil water.

D. major also exploits ant-deposited debris as a nitrogen source, and we used nitrogen isotope analysis to quantify the extent of this (the other likely nitrogen source, rainwater, has $\delta^{15}\text{N}$ values different from ant-provided debris). *Dischidia nummularia*, which does not possess ant leaves but grows on the same host trees as *D. major*, has access to the same nitrogen sources, with the exception of the ant-provided debris. We found that the mean $\delta^{15}\text{N}_{\text{plant}}$ value of *D. nummularia* was -3.5‰ ; in contrast, debris deposited in the ant leaves of *D. major* was significantly more enriched in ^{15}N ($U_{[5,2]}=0$, $P<0.05$; Fig. 3). This is expected, given that ant debris is composed mostly of scavenged insect parts⁴, and animal $\delta^{15}\text{N}$ values tend to be 3‰ greater than those of their food sources¹⁶. *D. major* leaves were significantly enriched in ^{15}N compared to those of *D. nummularia* ($U_{[17,2]}=2$, $P<0.05$; Fig. 3), suggesting that *D. major* absorbs nitrogen from ant-deposited debris. On average, *D. major* received about 29% of its nitrogen from ant debris (Fig. 3). Ant-related nitrogen could provide a large benefit to *D. major*, as nitrogen deficiency may be the major factor limiting epiphytic growth in the light-rich kerangas forests¹⁷.

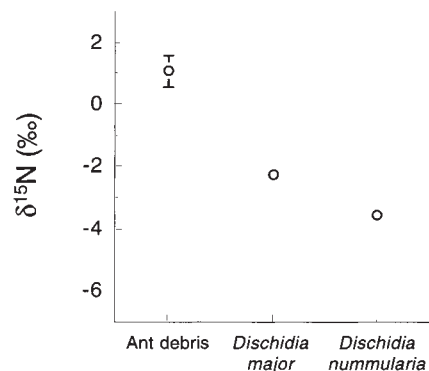


FIG. 3 Mean nitrogen isotope values (± 1 s.e.) for ant debris ($n=5$) and random leaves of *Dischidia major* ($n=17$) and *D. nummularia* ($n=2$). Experimental procedure is described in Fig. 2 legend. Standard error bars are not visible when the standard error is smaller than the plot symbol. For *D. major*, we calculated the per cent nitrogen content from ant debris as

$$\% \text{N from ant debris} = \frac{\delta^{15}\text{N}_{D, \text{major}} - \delta^{15}\text{N}_{D, \text{nummularia}}}{\delta^{15}\text{N}_{\text{ant debris}} - \delta^{15}\text{N}_{D, \text{nummularia}}} \times 100\%$$

where $\delta^{15}\text{N}_{D, \text{major}}$ and $\delta^{15}\text{N}_{D, \text{nummularia}}$ are the nitrogen isotope ratios of leaves of the two respective species. On average, *D. major* received 29% of its nitrogen from ant debris. (This estimate is conservative, as both *D. major* and *D. nummularia* often grow roots into ant carton (a mâché used for nest building) located on the host tree trunks, and the epiphyte may exploit this source as well¹⁷.)

Our observations strongly support Janzen's⁴ and Huxley's^{8,9} hypotheses of mutualism between *D. major* and *Philidris*. In exchange for shelter, ants provide significant amounts of two limiting resources: carbon dioxide and nitrogen. Both features could either expand the realized niche of *D. major*, enabling it to colonize hotter, drier habitats, or could provide *D. major* with a competitive edge over other epiphytes in this nutrient-poor ecosystem.

Finally, epiphytes in other tropical regions (including those of Central and South America, Papua New Guinea, the Philippines and Australia) have various structures occupied by ants^{4,8,9,17,23}. In a facultative myrmecophytic relationship involving an ant-occupied orchid from the neotropics, Fisher *et al.*²⁴ have used stable carbon isotopes to quantify the extent to which ants may forage on their own host plant. We may eventually be able to combine the two approaches and examine reciprocal benefits between plants and ants. □

Received 12 August 1994; accepted 28 March 1995.

1. Anderson, J. A. R. *Gdms Bull. Singapore* **20**, 131–228 (1963).
2. Brünig, E. F. W. O. *UNESCO Symp. Ecol. Res. Humid Trop. Veg., Kuching* **1963**, 289–313 (1965).
3. Ashton, P. S. *Malayan Nature J.* **24**, 151–162 (1971).
4. Janzen, D. H. *Biotropica* **6**, 237–259 (1974).

5. Shattuck, S. O. *Sociobiology* **21**, 1–181 (1992).
6. Kerr, A. F. G. Br. *Scient. Proc. R. Dublin Soc.* **13**, 293–309 (1912).
7. Griffith, W. *Trans. Linn. Soc. Lond.* **20**, 387–390 (1846).
8. Huxley, C. R. *Biol. Rev.* **55**, 321–340 (1980).
9. Huxley, C. R. in *Insects and the Plant Surface* (eds Juniper, B. & Southwood, R.) 255–282 (Arnold, London, 1986).
10. Winter, K., Wallace, B. J., Stocker, G. C. & Roksandic, Z. *Oecologia* **57**, 129–141 (1983).
11. Treseder, K. K. thesis, Univ. Utah (1994).
12. Farquhar, G. D., Ehleringer, J. R. & Hubick, K. T. A. *Rev. Pl. Physiol. molec. Biol.* **40**, 503–537 (1989).
13. Keeling, C. D., Mook, W. G. & Tans, P. P. *Nature* **277**, 121–123 (1979).
14. DeNiro, M. J. & Epstein, S. *Geol. Soc. Am. Abstr. Prog.* **8**, 834–835 (1976).
15. DeNiro, M. J. & Epstein, S. *Geochim. cosmochim. Acta* **42**, 495–506 (1978).
16. Schoeninger, M. J. & DeNiro, M. J. *Geochim. cosmochim. Acta* **48**, 625–639 (1984).
17. Davidson, D. W. & Epstein, W. W. in *Vascular Plants as Epiphytes* (ed. Lüttge, U.) 200–233 (Springer, Heidelberg, 1989).
18. Rico-Gray, V., Barber, J. T., Thien, L. B., Ellgaard, E. G. & Toney, J. J. *Am. J. Bot.* **76**, 603–608 (1989).
19. Benzing, D. H. *Bull. Torrey Bot.* **97**, 109–115 (1970).
20. Madison, M. *Selbyana* **5**, 107–115 (1979).
21. Gay, H. *Biol. J. Linn. Soc. Lond.* **50**, 221–233 (1993).
22. Huxley, C. R. *New Phytol.* **80**, 231–268 (1978).
23. Rickson, F. R. *Am. J. Bot.* **66**, 87–90 (1979).
24. Fisher, B. L., Sternberg, L. S. L. & Price, D. *Oecologia* **83**, 263–266 (1990).
25. Ehleringer, J. R. & Osmond, C. B. in *Plant Physiological Ecology* (eds Pearcy, R. W., Ehleringer, J. R., Mooney, H. A. & Rundel, P. W.) 281–300 (Chapman & Hall, London, 1991).

ACKNOWLEDGEMENTS. We thank D. R. Sandquist for advice, discussions and critical reading of the manuscript; B. E. Howlett for help in the field; E. Medina and T. G. Whitham for comments on the manuscript; and C. S. Cook, A. J. Leffler and C. F. Kitty for technical assistance. This work was supported by grants from the Explorer's Club, Sigma Xi, NSF, and the HHMI Undergraduate Research Program (University of Utah).

Visual motion aftereffect in human cortical area MT revealed by functional magnetic resonance imaging

Roger B. H. Tootell, John B. Reppas, Anders M. Dale^{*†}, Rodney B. Look, Martin I. Sereno^{*}, Rafael Malach, Thomas J. Brady & Bruce R. Rosen

Massachusetts General Hospital Nuclear Magnetic Resonance Center, 149 13th Street, Charlestown, Massachusetts 02129, USA

^{*} Department of Cognitive Sciences, University of California at San Diego, La Jolla, California 92093-0515, USA

[†] Department of Neurophysiology, University of Oslo, Oslo, Norway

FUNCTIONAL magnetic resonance imaging (fMRI)^{1–3} was used to measure local haemodynamic changes (reflecting electrical activity) in human visual cortex during production of the visual motion aftereffect, also known as the waterfall illusion^{4,5}. As in previous studies^{6–9}, human cortical area MT (V5) responded much better to moving than to stationary visual stimuli. Here we demonstrate a clear increase in activity in MT when subjects viewed a stationary stimulus undergoing illusory motion, following adaptation to stimuli moving in a single local direction. Control stimuli moving in reversing, opposed directions produced neither a perceptual motion aftereffect nor elevated fMRI levels postadaptation. The time course of the motion aftereffect (measured in parallel psychophysical tests) was essentially identical to the time course of the fMRI motion aftereffect. Because the motion aftereffect is direction specific, this indicates that cells in human area MT are also direction specific. In five other retinotopically defined cortical areas, similar motion-specific aftereffects were smaller than those in MT or absent.

Prolonged viewing of a stimulus moving in one direction (as when staring fixedly at a waterfall) makes stationary stimuli appear to move in the opposite direction immediately afterwards^{4,5}. It is not yet known which areas of human brain

TABLE 1 fMRI activation during the motion aftereffect

Cortical area	MAE excitation (%)	MAE motion selectivity (%)
MT	70	100
VP	29	53
V3a	26	48
V2	23	40
V3	~0	~0
V1	~0	~0

The maximum averaged difference in amplitude produced following single- minus reversing-direction rings was measured (as in Fig. 3) in MT and in five additional cortical visual areas. In the middle column, the amplitude of this fMRI MAE difference amplitude is expressed as a percentage of the averaged amplitude produced by the moving stimulus itself (moving minus initial fixation conditions), for each visual area. In the right column, the MAE amplitudes are expressed relative to the motion selectivity of each visual area (moving minus stationary conditions). The visual cortical areas were functionally labelled in the same scanning sessions in which the MAE data were taken, using retinotopically specific stimuli^{9,20–22}. Cortical area names are generalized from apparently homologous areas in macaque. Human area V1 (primary visual cortex) was defined as the area containing a complete representation of left or right hemifield buried within the calcarine fissure, extending superiorly and inferiorly along the medial bank. Area V2 is that area surrounding the vertical meridian representation of V1, containing a mirror-image representation of that same hemifield, split into an inferior and a superior region, located mostly along the medial bank. Area V3 was the mirror-symmetrical representation of the inferior visual quarter-field, bordering superior V2, extending from the horizontal meridian representation at the superior V2/V3 border to the vertical meridian representation at the V3/V3a border, representing only the inferior quarter-field. Human area V3 is proportionately wider in humans than in macaques. Area V3a borders V3 on its inferior/posterior aspect, and it again contains a complete representation of the left or right hemifield. Area VP is retinotopically mirror-symmetrical to area V3, bordering inferior V2, containing a quarter-field representation of the superior visual field. Area MT (V5) is a small oval area on the lateral surface near the occipito-temporal-parietal junction, which is motion selective and has very high contrast sensitivity⁹. Area MST can be distinguished from MT using certain tests involving eye movements and motion coherence, described elsewhere.

mediate this visual motion aftereffect (MAE), although the motion-specific area MT⁶⁻¹¹ (also known as V5) is one likely candidate. In non-human primates, most area MT neurons are both motion- and direction specific, and direction-specific cells have been presumed to underlie the motion aftereffect, either in monkey MT¹² or in non-primates^{13, 18}.

The stimuli used in this study were concentric rings (0.5 cycles deg⁻¹, duty cycle=0.2), either moving (7 deg s⁻¹) or stationary (see Fig. 1a). Human area MT was activated much more by moving than by stationary stimuli (see Figs 1b and 2), as described previously⁶⁻⁹. In addition, we found a clear increase in magnetic resonance (MR) signal amplitude during viewing of stationary stimuli when they were preceded by an adaptation stimulus moving continuously in a single direction. During these times the stationary stimuli appeared to be moving owing to a prominent motion aftereffect. Such an elevated MR signal did not occur following adaptation to otherwise identical stimuli that continually reversed direction (0.5 Hz), and neither did such stimuli induce a perceptual motion aftereffect. One obvious interpretation is that the signal remained high in area MT

because the stimulus appeared to be moving, even though it was in fact stationary. Because this MR increase is direction specific, and because it was accompanied by a perceptual motion aftereffect, we refer to it as a fMRI motion aftereffect.

The fMRI motion aftereffect was isolated by subtracting its signal from that of control conditions not showing a perceptual motion aftereffect (see Fig. 3). The peak amplitude of that averaged motion aftereffect reached 1.9%, equal to the mean steady-state difference in MR signal produced by rings that were actually moving versus those that were stationary.

If the fMRI motion aftereffect is related to the perceptual motion aftereffect, the time course of the two aftereffects is expected to be similar. To test this we showed subjects the MAE stimulus used in the magnet (e.g. expanding rings for 40 s, followed by stationary rings), with modifications to allow measurement of the psychophysical motion aftereffect (two-alternative forced choice motion-nulling staircase procedure¹⁹). The time course of the decay of fMRI and psychophysical motion aftereffects were very similar, with best-fit exponents of 8.3 s and 9.2 s, respectively (see Fig. 3).

FIG. 1 Stimulus used in this experiment, and human cortical visual area MT (V5) activated by that stimulus. **a**, A stationary view of the stimulus used in these experiments. **b**, A three-dimensional reconstruction of the brain of one subject (M.S.), derived entirely from magnetic resonance data. The brain is shown in both normal and 'inflated' format. The inflated version was made by relaxing curvature while approximately preserving local area and local angles²². Sulcal cortex (concave) is dark magenta and gyral cortex (convex) is lighter magenta. The fMRI activity produced by moving minus stationary rings is coded in a pseudocolour scale varying from saturated magenta (threshold) to white (maximum activity). The prominent white patch on the bottom right lateral surface is area MT (V5), a region that responds selectively in fMRI experiments to moving visual stimuli, as compared with stationary stimuli⁶⁻⁹. For greater functional specificity the stimulus used to produce **b** was low in luminance contrast⁹.

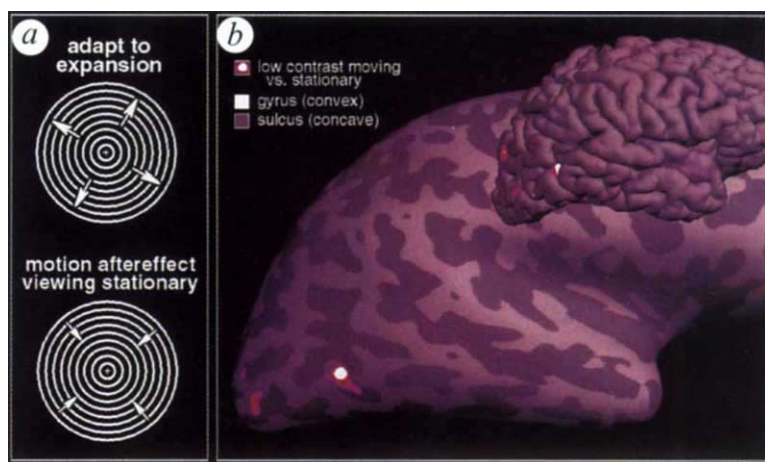
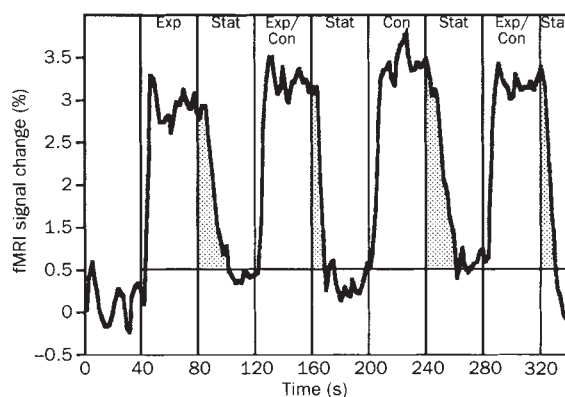
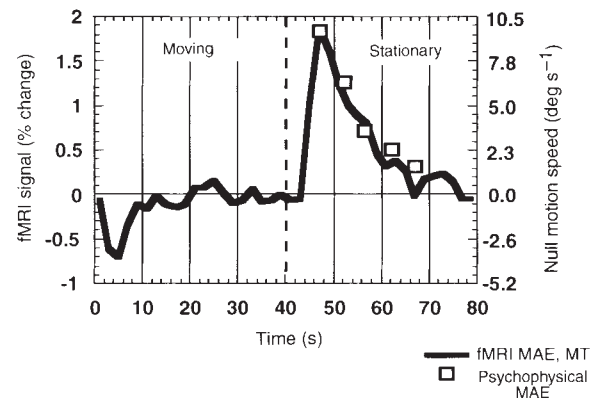


FIG. 2 Averaged MR time course showing responses during real and illusory visual motion. Subjects (12) were scanned in a 1.5 T GE MR scanner using echo-planar imaging (Advanced NMR) and a 5" radial surface coil positioned over visual cortex. Five or six slices (4–6 mm thick) were oriented approximately perpendicular to the calcarine fissure. In-plane resolution was 3.1 × 3.1 mm. In most subjects, head motion was greatly minimized by the use of a bite bar. We tested for motion artefact by presenting all images within a scan as a movie, or by analysis of difference images taken from different time periods. Any scans in which head motion within a scan exceeded approximately 2 mm were discarded. Asymmetric spin echo sequences (repetition time, 2,000 ms, echo time, 80 ms, 180° refocusing pulse offset by -25 ms) were used to measure 'activation' (local increases in blood flow and oxygenation). Subjects were shown a pattern of concentric rings (50° diameter) surrounding a fixation spot, with 1 image every 2 s. In different periods (typically 40" long) during each scan (5' 40"), the moving rings were either (1) continuously expanding (Exp), (2) continuously contracting (Con), or (3) reversing direction (expanding or contracting, Exp/Con) at 0.5 Hz. Radially symmetric stimuli (rings) were used to induce the motion aftereffect because such stimuli do not induce optokinetic nystagmus, a potential confound. However, within each local visual region, the direction of stimulus movement was either unidirectional (conditions 1 and 2) or bidirectional (condition 3). Following every period of moving rings, stationary rings, otherwise equal (Stat), were presented. Following the periods of continuous unidirectional local motion (the expanding or contracting stimuli), a profound visual motion aftereffect was seen in the (physically stationary) rings. Following the periods of reversing direction, no motion aftereffect was reported. The averaged MR signal from all 38 scans in human area MT/V5 is shown. The MR response amplitudes to both the single- and reversing-direction



stimuli were approximately equal in amplitude. The response to stationary stimuli was only slightly greater than that to a blank field. However, the fMRI response immediately following the single-direction stimuli (i.e. when the motion aftereffect was visible) remained high for some time after stimulus offset, significantly longer than that predicted by the normal temporal response of the fMRI signal^{1,3,27}. MR responses following reversing-direction stimuli (when motion aftereffects were absent) instead returned promptly to the steady-state response level produced by stationary stimuli. One interpretation is that area MT remained activated because the stimulus appeared to be moving, even though it was physically stationary.

FIG. 3 Isolated fMRI motion aftereffect and its relation to psychophysics. This time course shows MR amplitudes during and after single-direction conditions, minus amplitudes during and after reversing-direction conditions. All data are included. During the first 40 s (20 images), the figure indicates no steady-state difference between the activation produced by single-direction versus reversing-direction stimuli. Thus the subsequent fMRI aftereffect cannot be attributed to unequal levels of activation. Thereafter the subjects viewed stationary stimuli, but the MR signal remained high following single-direction stimuli (relative to reversing-direction stimuli) for approximately 20 s. The time course of the psychophysical motion aftereffect was measured in response to the fMRI stimuli, with modifications. As in the fMRI experiments, subjects viewed rings moving in a single local direction (e.g. expanding) for 40 s, followed by stationary rings. After a delay of either 0, 4.5, 9.5, 14.5 or 19.5 s, the stationary rings were replaced by a blank field of uniform grey for 0.5 s. This was followed by a pattern of sinusoidal rings (same spatial frequency) of low (~5%) contrast, presented for 0.5 s. The subjects then indicated (using a two-alternative forced-choice procedure) which (illusory) direction the rings were moving and the speed and direction of the rings was adjusted opposite to the indicated direction in small increments at every button press. Thus the ring speed converged to a velocity and direction equal but opposite to that of the illusory motion, independently for every delay measured (0, 5, 10, 15 and 20 s). Data from all four subjects were averaged. The average standard deviation at plateau was 0.37 deg s^{-1} . The psychophysical data are shown as five squares. To facilitate comparisons between fMRI and psychophysical motion aftereffects, the psychophysical time course was shifted to the right by 7 s to compensate for the known delay in the fMRI haemodynamics^{1,3,27}, and peak fMRI and psychophysical amplitudes were normalized on different y-axes (psychophysical, right). There is an excellent fit between the time courses of the psychophysical and the fMRI motion aftereffects. Previous models of the motion aftereffect predict a downward deflection of fMRI signal rather than the positive inflection obtained. Single-unit studies in non-primate species, from areas other than MT, suggest that direction-specific cells tuned to the adapting stimulus direction habituate (decrease firing) during and after



the adapting period¹³⁻¹⁸. Thus the population firing of cells preferring opposing directions would predominate in response to a stationary stimulus after single-motion adaptation, leading to an illusory motion in the opposing direction, and perhaps a decrease in MR amplitude. However, other evidence and a different model are consistent with the present fMRI findings. In the only electrophysiological study in monkey area MT¹² on the motion aftereffect, cells preferring directions opposite to the adapting direction increased firing postadaptation. As in other studies, cells preferring the adaptation direction also decreased their firing rate. Together, these indicate that cells tuned to opposing directions tonically inhibit each other. Tonic, reciprocal, direction-specific inhibition of monkey MT cells is suggested by other evidence as well²⁸⁻³⁰. Because cortical single units typically have low baseline rates, this model (and our fMRI data) suggests that poststimulus excitation overcomes poststimulus inhibition, because inhibition cannot decrease below zero spikes per second, whereas excitation is not so limited. However, because the physiological basis underlying fMRI is incompletely understood, such a model is obviously speculative.

Presumptive MT (V5) has been shown to be motion selective⁶⁻⁹ (see Fig. 1b). However, the prominent feature of single units in monkey MT is that they are direction specific as well as motion specific. To our knowledge, the direction-specific motion aftereffect revealed here is the first physiological evidence for direction specificity in human cortical visual area MT. This is important because the human area chosen to be 'MT' or 'V5' is only one of several motion-selective cortical areas⁷⁻⁹, and many visual neurons lacking direction specificity will respond better to moving than to stationary stimuli.

The high spatial resolution of the fMRI technique, coupled with retinotopically specific visual stimulation, made it possible to localize accurately other cortical visual areas^{9,20-22} during the same experiments (see Table 1). This allowed us to investigate which of those other areas, if any, also show a fMRI motion aftereffect, and to measure its relative amplitude.

These data are shown in Table 1. As expected, area MT showed the highest MAE activation ratio of all the measured areas. Smaller fMRI motion aftereffects were seen in areas V2 and V3a. This is compatible with the presence of a minority percentage of direction-selective single units in V2 and V3a, and with reciprocal connections between MT, V2 and V3a in monkey (see ref. 23). A robust MAE appeared to be present in presumptive human MST as well, but more data are required.

Because the perceptual effect depends logically on the presence of direction specificity, a trivial explanation for our results is that the amplitude of the direction-specific fMRI MAE is simply proportional to the motion specificity of the area. However, when this correction was applied (Table 1, middle column), the MAE values still differed markedly across the different cortical areas. The fMRI motion aftereffect was largest in MT even when corrected for the greater motion selectivity of that area.

These results suggest that MT and certain areas 'upstream' from it are selectively activated during the perceptual motion aftereffect. Other stimuli, not tested here, produce a motion aftereffect which (based on the extent of binocular transfer)

appears to tap higher stages of motion processing than those tested here^{24,25}; such stimuli might produce a correspondingly more specific fMRI localization as well. The illusory motion reported here may also be related to other reports of illusory motion²⁶ or illusory lack of motion at equiluminance⁹. □

Received 19 December 1994; accepted 18 April 1995.

- Kwong, K. K. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 5675-5679 (1992).
- Ogawa, S. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 5951-5955 (1992).
- Bandettini, P. A., Wong, E. C., Hinks, R. S., Tikofsky, R. S. & Hyde, J. S. *Magn. Reson. Med.* **25**, 390-397 (1992).
- Adams, R. Phil. Mag. **5**, 373-374 (1834).
- Wohlgemuth, A. Br. J. Psychol. (Suppl.) **1**, 1-117 (1911).
- Zeki, S. et al. *J. Neurosci.* **11**, 641-649 (1991).
- Watson, J. D. G. et al. *Cerebr. Cort.* **3**, 79-94 (1993).
- Dupont, P., Orban, G. A., DeBruyn, B., Verbruggen, A. & Mortelmans, L. J. *Neurophysiol.* **72**, 1420-1424 (1994).
- Tootell, R. B. H. et al. *J. Neurosci.* **15**, 3215-3230 (1995).
- Clarke, S. & Miklosy, J. J. *comp. Neural.* **298**, 188-214 (1990).
- Tootell, R. B. H. & Taylor, J. B. *Cerebr. Cort.* **5**, 39-55 (1995).
- Petersen, S. E., Baker, J. F. & Allman, J. M. *Brain Res.* **346**, 146-150 (1985).
- Barlow, H. B. & Hill, R. M. *Nature* **200**, 1345-1347 (1963).
- Hammond, P., Mouat, G. S. V. & Smith, A. T. *Expl Brain Res.* **60**, 411-416 (1985).
- Marlin, S. G., Sohail, S. J., Hasan, J. & Cynader, M. S. *J. Neurophysiol.* **59**, 1314-1330 (1988).
- Movshon, J. A. & Lennie, P. *Nature* **278**, 850-852 (1979).
- Vautin, R. G. & Berkeley, M. A. *J. Neurophysiol.* **40**, 1051-1065 (1977).
- Von der Heydt, R., Hanny, P. & Adorjani, C. *Archs ital. Biol.* **116**, 248-254 (1978).
- Carney, T. & Shadlen, M. N. *Vision Res.* **33**, 1977-1995 (1993).
- Schneider, W., Noll, D. C. & Cohen, J. D. *Nature* **365**, 150-153 (1993).
- DeYoe, E. A. et al. *J. Neurosci. Meth.* **54**, 171-187 (1995).
- Sereno, M. I. et al. *Science* (in the press).
- Felleman, D. J. & Van Essen, D. C. *Cerebr. Cort.* **1**, 1-47 (1991).
- Nishida, S., Ashida, H. & Sato, T. *Vision Res.* **34**, 2707-2716 (1994).
- Nishida, S. & Sato, T. *Vision Res.* **32**, 1635-1646 (1992).
- Zeki, S., Watson, J. D. & Frackowiak, R. S. *Proc. R. Soc. B252*, 215-222 (1993).
- Friston, K. J., Jezzard, P. & Turner, R. *Hum. Brain Mapping* **1**, 153-171 (1994).
- Born, R. T. & Tootell, R. B. H. *Nature* **357**, 497-499 (1992).
- Allman, J. M., Baker, J. F., Newsome, W. T. & Petersen, S. E. in *Cortical Sensory Organization Vol. 2. Multiple Visual Areas* (ed. Woolsey, C. N.) 171-185 (Humana, Clifton, NJ, 1981).
- Lagae, L., Gulyas, B., Raiguel, S. & Orban, G. A. *Brain Res.* **496**, 361-367 (1989).

ACKNOWLEDGEMENTS. This study was supported by an NEI grant to R.B.H.T. We thank several subjects, including L. Kobierski, W. Kennedy, W. Vanduffel and H. Aronen. We also thank T. Campbell, who ran the MR scanner on some of the scans, and many others in the MGH NMR Center.

Sequential signalling during *Caenorhabditis elegans* vulval induction

Jeffrey S. Simske & Stuart K. Kim

Department of Developmental Biology, Stanford University Medical School, Stanford, California 94305, USA

DURING the induction of the *Caenorhabditis elegans* vulva, cell signalling causes initially equipotent cells to express a reproducible pattern of cell fates^{1,2}. The position of the anchor cell determines the pattern of vulval precursor cell fates, such that the closest precursor cell (P6.p) expresses the primary cell fate, the next closest cells (P5.p and P7.p) both express the secondary cell fate, and each of the precursor cells located at a distance (P3.p, P4.p and P8.p) express the tertiary cell fate (Fig. 1a)³⁻⁵. We present data indicating that this stereotypical pattern of cell fates can be generated by sequential signals. We identified genetic mosaic animals in which P5.p and P7.p were defective in the anchor-cell signal-transduction pathway and observed that these cells adopted the secondary cell fate, indicating that anchor-cell signal transduction is not required for the expression of the secondary cell fate. These results suggest that the anchor cell induces P6.p to express the primary cell fate, and that P6.p subsequently induces P5.p and P7.p to express the secondary cell fate.

Several models could explain how the wild-type pattern of vulval cell fates is established by the anchor cell. For instance, the anchor cell could generate a spatially graded signal, encoded by *lin-3* (ref. 6), such that high, medium or low signal concentrations could cause vulval precursor cells to express the primary, secondary or tertiary cell fates, respectively^{4,5}. Alternatively, the anchor cell could initiate a signalling cascade, such that the anchor cell induces P6.p to express the primary cell fate, and P6.p subsequently induces P5.p and P7.p to express the secondary cell fate.

To distinguish between the graded and sequential signalling models, we used mosaic analysis to identify animals in which the anchor-cell signalling pathway was normal in P6.p but defective in P5.p and P7.p (because of mutations in either *let-23 receptor* or *lin-7*). LET-23 is likely to be the receptor for the anchor-cell signal; it is a receptor tyrosine kinase⁷ that activates a highly conserved signal transduction pathway, including homologues of GRB2, Ras, Raf and MAP kinase (reviewed in ref. 2). Furthermore, *let-23* mutants have no vulva because each Pn.p cell expresses the uninduced (tertiary) cell fate instead of induced (primary and secondary) cell fates⁸ (Table 1 row 2, and Table 2a row 2); *lin-7* mutants also have no vulva, because of an anchor-cell signalling defect in the Pn.p cells⁸⁻¹⁰ (J.S.S. and S.K.K., unpublished observations).

The two models make different predictions regarding the pattern of cell fates expressed by these *let-23* or *lin-7* mosaic animals. The graded-signal model predicts that P5.p and P7.p will express the tertiary cell fate, as anchor-cell signal transduction would be required in P5.p and P7.p for them to express the induced, secondary cell fate (Fig. 1b, left). The sequential signalling model predicts that P5.p and P7.p will express the secondary cell fate, as a second (lateral) signal from P6.p, rather than the anchor-cell signal, would specify the secondary cell fate (Fig. 1b, right).

To generate *let-23* mosaic animals, we constructed a strain in which wild-type *let-23* and several cell lineage marker genes (including *ncl-1*) were expressed from an extrachromosomal array. We used the *ncl-1* marker to identify animals that had spontaneously lost the extrachromosomal array during development¹¹⁻¹³, and then scored the cell fates expressed by the vulval precursor cells. A similar approach was used for *lin-7* mosaic analysis. As anticipated, we found that *let-23* acts in the vulval precursor cells to transduce the anchor-cell signal. Mosaic

animals lacked a vulva if wild-type *let-23* activity was missing in all of the vulval precursor cells, whereas animals almost always had a normal vulva if the array was present in the vulval precursor cells (Fig. 2a).

We also identified mosaic animals in which alternate vulval precursor cells had lost the extrachromosomal array (Tables 1 and 2). We first performed this experiment using *sy1*, a weak *let-23* allele¹⁴ (Table 2(a)). We then repeated the experiment using *sy15*, a *let-23* null allele¹⁴, to show that the results obtained using *sy1* (see below) did not merely reflect residual *let-23* activity resulting from a weak mutation (Tables 1 and 2b). Finally, we tested a second gene involved in anchor-cell signal transduction (*lin-7*) to show that our results are not only observed using *let-23* (Table 2c). We scored the cell fates expressed by the vulval precursor cells in *let-23(sy15)* mosaic animals by directly observing the Pn.p cell lineages (Table 1). We also inferred the cell fates expressed in all three classes of mosaic animals by scoring the number, position and morphology of the vulval cells in the fourth larval stage, and by scoring the function of the vulva in the adult (Table 2). We obtained similar results using each genotype and scoring method.

First, we found that P6.p only adopted the primary cell fate if it expressed wild-type *let-23* or *lin-7* (Tables 1 and 2), indicating that *let-23* and *lin-7* normally act in P6.p to transduce the anchor-cell signal.

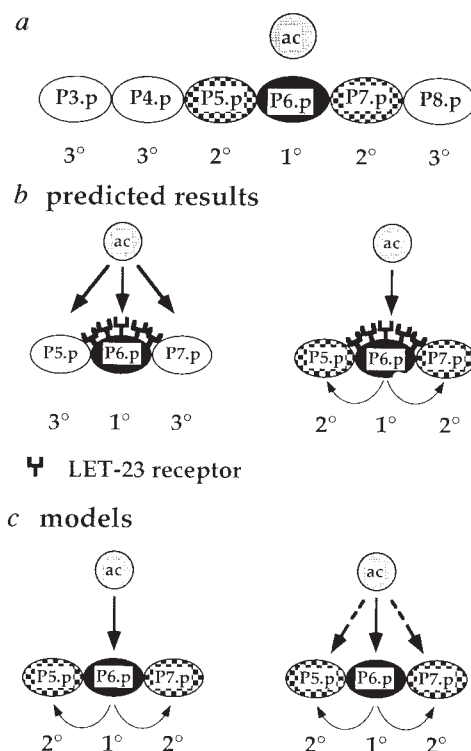


FIG. 1 a, Organization of cells involved in vulval induction. The vulval precursor cells (P3.p to P8.p) are arranged in an anterior to posterior row along the ventral midline. The gonadal anchor cell (ac) is dorsal to P6.p. b, Predicted cell fates expressed in mosaic animals lacking wild-type *let-23* or *lin-7* activity in P5.p and P7.p according to the graded and sequential signalling models. c, Sequential and combined signalling models. A black arrow indicates the anchor cell (ac) signal. Curved arrows indicate the lateral signal. Left, an inductive signal from the anchor cell induces the expression of the primary (1°) cell fate by P6.p. P6.p subsequently induces the expression of the secondary (2°) cell fate in adjacent Pn.p cells. Right, an inductive signal from P6.p as well as medium levels of anchor-cell signal (dashed black arrows) may induce the secondary cell fate in P5.p and P7.p.

TABLE 1 Vulval cell lineages in *let-23(sy15)* mosaic animals

Row	Loss at	P3.p	P4.p	P5.p	P6.p	P7.p	P8.p	Vulval phenotype
1	WT	S or SS	SS	LLTN	TTTT	NTLL	SS	normal
2	<i>let-23</i>	S or SS	SS	SS	SS	SS	SS	Vul
3	AB.pla	<i>SS</i>	SS	<i>LLTN</i>	TTTT	<i>NTLL</i>	SS	normal
4	AB.pla	S	<i>SS</i>	<i>LLTN</i>	TTTT	<i>NTLL</i>	SS	Pvl
5	AB.plap	S	<i>SS</i>	<i>LLTN</i>	TTTT	NTLL	<i>SS</i>	normal
6	AB.pr	S	SS	<i>LLTN</i>	TOTT	NTLL	<i>SS</i>	NT
7	AB.pl/rappap	SS	SS	LLON	TTOT	<i>NTLL</i>	SS	Egl
8	AB.plap	S	<i>LLTN</i>	TTTT	<i>NTLL</i>	SS	<i>SS</i>	Egl
9	AB.pla	S	<i>LLTN</i>	TTTT	<i>NTLL</i>	<i>SS</i>	SS	normal
10	AB.plap	<i>SS</i>	SLLO	OTTT	<i>NTLL</i>	<i>SS</i>	SS	NT
11	AB.plap	<i>SS</i>	SOT	TTTT	<i>NTLL</i>	<i>SS</i>	SS	normal
12	AB.pr	<i>SS</i>	LLON	OOOO	<i>NTLL</i>	TOOT	<i>SS</i>	NT
13	AB.pl/rapp	<i>SS</i>	SS	SS	<i>SS</i>	SS	<i>SS</i>	Vul
14	AB.pl/rapp	<i>SS</i>	SS	SS	<i>SS</i>	SS	<i>SS</i>	Vul
15	AB.prap	<i>SS</i>	SS	SS	<i>SS</i>	<i>SS</i>	SS	Vul

Pn.p cell lineages expressed in *let-23(sy15)* mosaic animals. Bold indicates Pn.p progeny cells that were non-Ncl (contained the array), script indicates Pn.p progeny cells that were Ncl (lacked the array), and italics indicates cells in which the *ncl-1* phenotype could not be scored. Cells are non-dividing (N), longitudinally dividing (L), transversely dividing (T), dividing along an oblique axis (O) or joining the hypodermal syncytium (S). Underlining indicates that a cell remained attached to the cuticle. Thick- and thin-lined boxes indicate primary and secondary cell fates, respectively. S or SS denotes a tertiary cell fate. Mosaic animals were identified at the L2 stage from approximately 1200 animals, and were subsequently mounted on 5% agar pads for observation using Nomarski optics to follow the Pn.p cell lineages through the final round of division. Cell fates were determined by scoring the pattern of cell divisions and progeny cell types produced by each Pn.p cell^{4,22}, and the phenotype of the vulva in the adult, because partial induction or an abnormal cell fate pattern usually results in the formation of an abnormal vulva²³. It was noted whether animals had a protruding vulva (Pvl), were swollen with eggs (Egl) or did not lay eggs (Vul); NT, not tested. The ventral row of six vulval precursor cells is formed when three cells from the right side of the animal (derived from AB.pr) interdigitate with three cells from the left side (derived from AB.pl). Thus loss of the extrachromosomal array in either AB.pl or AB.pr (or from progeny of these cells) results in mosaic animals in which only one cell out of each pair of Pn.p cells (P3.p/P4.p, P5.p/P6.p, P7.p/P8.p) carries the array. In these mosaic animals, the anchor cell usually migrates to a position dorsal to the Pn.p cell that adopts the primary cell fate. In mosaic animals in which P6.p lacks wild-type *let-23*, P5.p usually expresses the primary cell fate. This suggests that, in the wild type, P5.p expresses the secondary fate either because LET-23 expressed by P6.p may sequester most of the anchor-cell signal and prevent anchor-cell signal from inducing P5.p, or because the lateral signal from P6.p may be stronger than the anchor-cell signal. Induction of the secondary cell fate in cells lacking the array is unlikely to be due to perdurance of *let-23* RNA or protein expressed before loss of the array, as *let-23* reporter gene expression was not detected before array loss (J.S.S. and S.K.K., unpublished observations). For *let-23(sy15)* mosaic analysis, *unc-29(e1072)*; *let-23(sy15)*, *unc-4(e120)*; *ncl-1(e1942)* animals were transformed as for *let-23(sy1)* (Fig. 2 legend), except that *let-23* plasmid pk7-13.8 was injected at 100 $\mu\text{g ml}^{-1}$ and *unc-30* DNA was replaced with *unc-4* DNA (plasmid pNC4-21)²⁴ (D. Miller, personal communication), generating the strain SD314. The strongest *let-23* allele is *let-23(sy15)*, which behaves as a null in gene-dosage experiments¹⁴.

Second, in 96% ($n=89$) of the cases observed, we found that a Pn.p cell that could not transduce the anchor-cell signal (because it lacked wild-type *let-23* or *lin-7*) expressed the secondary cell fate when adjacent to a primary cell (Tables 1 and 2, Fig. 2b, c). For example, in the mosaic animal described in Table 1 row 3, P6.p expressed wild-type *let-23* and thus adopted the primary cell fate. P5.p and P7.p both lacked wild-type *let-23* and should not be able to receive directly the anchor-cell signal. However, both cells expressed the secondary cell fate, indicating that anchor-cell signal transduction is not required in P5.p and P7.p for the expression of the secondary cell fate. These results suggest that a lateral signal from P6.p can induce expression of the secondary cell fate independently of the anchor-cell signal.

Third, in nine mosaic animals, we observed cell fate patterns that were inconsistent with a simple graded-signal model (Table 1 row 8; Table 2a, rows 12 and 13; Table 2b, row 9; Table 2c, row 11). In these animals, P7.p should be able to respond to the postulated graded anchor-cell signal, as it expresses wild-type *let-23* or *lin-7*. However, P7.p should not receive the lateral signal, because P6.p did not express the primary cell fate. In all nine cases, P7.p adopted the tertiary rather than the secondary cell

fate. These results suggest that a graded anchor-cell signal is usually not sufficient to induce expression of the secondary cell fate by P7.p in the absence of the lateral signal from P6.p.

Fourth, our results extend previous work on the lateral signal¹⁵ by showing that this signal can induce the expression of the secondary cell fate, rather than merely inhibit the expression of the primary cell fate. For example, in mosaic animals in which P5.p expressed the primary cell fate, P4.p usually expressed the secondary cell fate (Tables 1 and 2). P4.p is unlikely to receive inductive levels of the anchor-cell signal (because it is far from the anchor cell), so the expression of the secondary cell fate in these mosaic animals is probably due to induction by the lateral signal from P5.p.

In summary, our mosaic analyses suggest that the anchor-cell signal induces the closest Pn.p cell (P6.p) to express the primary cell fate via the *let-23*-mediated signalling pathway, and that the lateral signal from the primary cell then induces adjacent Pn.p cells (P5.p and P7.p) to express the secondary cell fate (Fig. 1c, left). The lateral signal is probably mediated by *lin-12* (reviewed in ref. 16), which is required for the expression of the secondary vulval cell fate and encodes a protein homologous to receptor proteins such as *Drosophila* Notch and *C. elegans* GLP-1.

TABLE 2 Vulval cell fates in mosaic animals

Row	Loss at	P3.p	P4.p	P5.p	P6.p	P7.p	P8.p	N	Cell number	Position/morphology	Vulval phenotype
(a) <i>let-23(sy1)</i> mosaic											
1	WT	3°	3°	2°	1°	2°	3°	many	22	normal	normal
2	<i>let-23(sy1)</i>	3°	3°	3°	3°	3°	3°	many	0	NA	Vul
3	AB.p	3°	3°	3°	3°	3°	3°	10	0	NA	Vul
4	AB.pr	3°	3°	2°	1°	2°	3°	1	22	normal	weak Egl
5	AB.pl/r	3°	3°	2°	1°	2°	3°	1	22	normal	normal
6	AB.pl/r	3°	3°	2°	1°	2°	3°	3	22	normal	normal
7	AB.pl	3°	3°	2°	1°	2°	3°	1	22	normal	normal
8	AB.plap	3°	3°	2°	1°	2°	3°	1	22	normal	normal
9	AB.pr	3°	2°	2°	1°*	2°	3°	1	29	abn P6.p	Muv
10	AB.pl	3°	2°	1°	2°	3°	3°	1	22	normal	normal
11	AB.pl	3°	2°	1°	2°	3°	3°	1	22	normal	normal
12	AB.pl/r	3°	2°	1°	2°	3°	3°	3	22	skewed	normal
13	AB.pr	3°	2°	1°	~2°†	3°	3°	1	18	abn P6.p	normal
14	AB.plap	3°	2°	1°	~2°*	1°	1°	1	>30	abn‡	Muv
(b) <i>let-23(sy15)</i> mosaic											
1	AB.pr	3°	3°	2°	1°	2°	3°	1	22	normal	normal
2	AB.pr	3°	3°	2°	1°	2°	3°	1	22	normal	normal
3	AB.pl/rapp	3°	3°	2°	1°	2°	3°	1	22	normal	normal
4	AB.pl/rappap	3°	3°	2°	1°	2°	3°	2	22	normal	Egl
5	AB.pr	3°	3°	~2°§	1°	2°	3°	1	20	abn P5.p	Vul
6	AB.pr	3°	3°	~2°	1°	2°	3°	1	21	abn P5.p	NT
7	AB.pr	3°	2°	1°	2°	3°	3°	1	22	normal	normal
8	AB.pl/rapp	3°	2°	1°	2°	3°	3°	1	22	normal	normal
9	AB.pl/rapp	3°	~2°¶	1°	2°	3°	3°	1	19	abn P4.p	weak Egl
10	AB.pra	3°	~2°#	1°	2°	3°	3°	1	21	abn P4.p	weak Egl
(c) <i>lin-7(e1413)</i> mosaic											
1	AB.pr	3°	3°	2°	1°	2°	3°	2	22	normal	normal
2	AB.pl/r	3°	3°	2°	1°	2°	3°	4	22	normal	normal
3	AB.pl/r	3°	3°	2°	1°	2°	3°	2	22	normal	normal
4	AB.pr	3°	3°	2°	1°	2°	3°	1	22	normal	normal
5	AB.pl	3°	3°	2°	1°	2°	3°	1	22	normal	normal
6	AB.pl/r	3°	3°	2°	1°	2°	3°	3	22	normal	normal
7	AB.pl	3°	3°	2°	1°	2°	3°	1	22	normal	Pvl
8	AB.pr	3°	3°	3°	1°	2°	3°	1	15	abn*	Vul
9	AB.pr	3°	3°	2°	1°	3°	3°	1	15	abn*	Vul
10	AB.pr	3°	3°	2°	1°	3°	3°	1	15	abn*	Egl
11	AB.pl/r	3°	2°	1°	2°	3°	3°	3	22	skewed	Egl
12	AB.pl	3°	2°	1°	2°	3°	3°	1	22	normal	normal
13	AB.pl	3°	2°	1°	2°	3°	3°	1	22	normal	normal
14	AB.pr	3°	2°	1°	2°	1	1°	1	>30	abn**	Muv
15	AB.pl	3°	3°	3°	2°	1°	3°	1	15	abn*	Egl/Pvl

Pattern of Pn.p cell fates expressed in (a) *let-23(sy1)*, (b) *let-23(sy15)* and (c) *lin-7(e1413)* mosaic animals. Presence of the array is indicated as in Table 1. Solid bars mark Pn.p cells that expressed induced cell fates (1°, primary; 2°, secondary). ‡ denotes an induced cell fate that could not be scored as either primary or secondary, ~2° denotes an incomplete secondary cell fate. The vulval precursor cell fates (1°, 2°, 3°) were determined for each of the mosaic animals by scoring the number of cell nuclei (cell number) that make up the vulva at the L4 larval stage, the final position and morphology of the Pn.p cell descendants. Skewed denotes animals in which the L4 vulval structure was bent posteriorly to interact with the anchor cell, positioned dorsal to P6.p descendants. The phenotype of the vulva in the adult was scored as in Table 1 with the following addition: Muv denotes animals that displayed a multi-vulval phenotype. For *lin-7* mosaic analysis, *unc-29(e1072)*; *lin-7(e1413)*; *ncl-1(e1942)*; *unc-30(e191)* animals were transformed as above, except that DNA containing a wild-type copy of *lin-7* (plasmid spe-5.0) (J.S.S. and S.K.K., unpublished observations) was used, generating the strain SD274. Mosaics animals were identified at the L2 to L4 stage from approximately 6,000, 2,000 and 10,000 animals for the *let-23(sy1)*, *let-23(sy15)* and *lin-7* mosaic analyses, respectively. The extrachromosomal arrays used in these mosaic analyses rescued the Vul phenotype of non-mosaic *let-23(sy1)*, *let-23(sy15)* and *lin-7(e1413)* animals with an efficiency of 78% (strain SD313; $n = 27$), 94% (strain SD314; $n = 18$) and 89% (strain SD274; $n = 18$), respectively. A fraction (17%, $n = 63$) of the primary cells did not induce both adjacent Pn.p cells to express the secondary cell fate (Table 1 row 10–12; Table 2a rows 13 and 14; Table 2b rows 5, 6, 9 and 10; Table 2c rows 8–10 and 15), suggesting that occasional weak expression of *let-23* or *lin-7* from the extrachromosomal array can sometimes result in poor expression of the inductive lateral signal. *lin-7(e1413)* is a stop mutation early in the *lin-7* coding region (J.S.S. and S.K.K., unpublished observations).

* The descendants of this precursor cell were disorganized at the L4 stage.

† NNN.

‡ Five precursor cells expressed induced fates resulting in excess vulval cells.

§ LLN.

|| LLNN.

¶ NTN.

NLTN.

* The vulval cells produced by the two induced precursor cells developed normally.

** Five precursor cells expressed induced fates resulting in excess vulval cells.

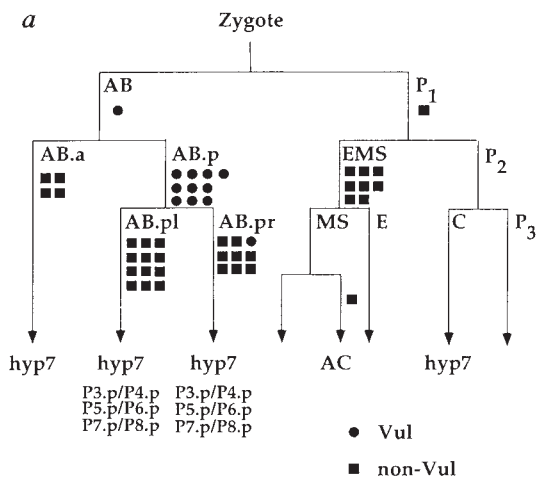
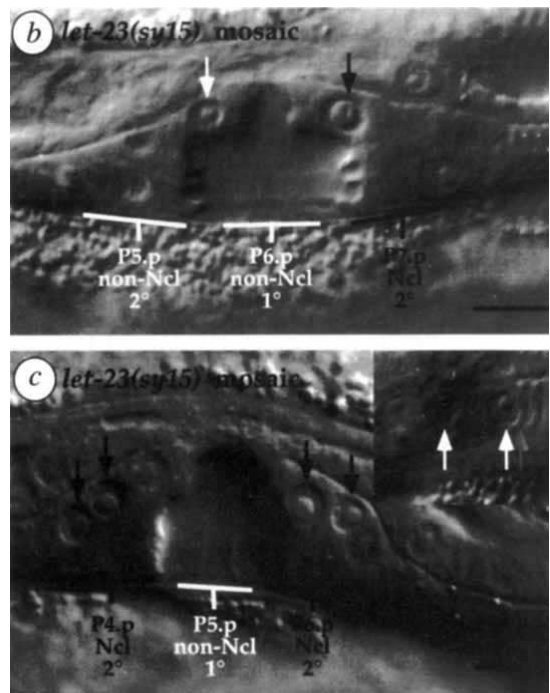


FIG. 2 Analysis of *let-23* mosaics. *a*, The pattern of cell divisions that generate cells known to be involved in vulval development is shown^{22,25}. Vertical lines represent single cells, and horizontal lines represent a cell division. Vertical arrows indicate that many progeny cells are generated. Circles and squares denote individual mosaic animals with a Vul or non-Vul phenotype, respectively. The position of the marker represents the cell division in which the array was lost based on scoring the *Ncl* phenotype of the cells in the mosaic animal. A loss at EMS could not be distinguished from a loss in MS, as the *ncl-1* phenotype could not be scored in descendants of the E lineage. AB.pr losses could not always be distinguished from AB.pra losses. Eleven *let-23(sy1)* mosaic animals were vulvaless when the array was lost in all vulval precursor cells, including one that lost the array in AB and 10 that lost the array in AB.p. Conversely, 14 mosaic animals had a normal vulva when wild-type *let-23* activity was present in all 6 vulval precursor cells but absent in other cells, including 4 in which most of the hyp-7 hypodermal syncytium lacked the array because of a loss in AB.a, and 10 in which the anchor cell lacked the array because of a loss in P₁, EMS or an MS daughter. Similar results have been observed by others (M. Koga and Y. Ohshima, personal communication). *b, c*, Nomarski photomicrographs of *let-23(sy15)* mosaic animals. Tissues (bars) or cells (arrows) derived from *Ncl* and non-*Ncl* Pn.p cells are denoted with black and white, respectively. *b*, Mosaic animal described in Table 2b row 4 shown in the late fourth larval stage. P5.p, P6.p and P7.p expressed secondary, primary and secondary cell fates, respectively, and generated a wild-type vulva. P7.p lacked the array because all of the P7.p descendants, including P7.paa (black arrow), were *Ncl*. Both P5.p and P6.p contained the array because all of their descendants, including P5.ppp (white arrow), were non-*Ncl*. *c*, Mosaic animal described in Table 1 row 9, shown in the early fourth larval stage. P4.p, P5.p and P6.p expressed secondary, primary and secondary cell fates, respectively. P4.p and P6.p lacked the array because their descendants (black arrows) were *Ncl*.



Inset: different focal plane of the same animal. P5.p contained the array because its descendants, including P5.papl and P5.ppal (white arrows), were non-*Ncl*. Scale bar, 10 μ m.

METHODS. *unc-29(e1072)*; *let-23(sy-1)*; *ncl-1(e1942)*; *unc-30(e191)* animals were transformed with DNAs (50 μ g ml⁻¹) containing wild-type copies of *unc-29* (cosmid C45D10), *let-23* (plasmid pk7-13.8), *ncl-1* (cosmid C33C3) and *unc-30* (cosmids C37G8 and G33F2), using the procedure of Fire²⁶ as modified by Mello²⁷, generating the strain SD313. Body muscle cells derived from the MS, C and D blastomeres require *unc-29* activity (L. Miller, personal communication); *unc-30* is required in the VD and DD neurons, which are derived from AB.p²⁸; and *ncl-1* is required for the normal nuclear morphology of most cells²⁹. The point in the cell lineage at which the array was lost was determined as previously described^{12,13,29}. Spontaneous loss of the extrachromosomal array generates a clone of cells that lack the array and thus lack wild-type *let-23* and *ncl-1* gene activities. Because the cell lineage is essentially invariant, the founder cell from which all *Ncl* progeny cells are derived can be inferred by scoring the *ncl-1* phenotype of progeny cells. Some mosaic animals were first identified because they expressed one marker (*unc-30*) but not the other (*unc-29*), or vice versa. The presence of the array in the germ-line was determined by showing whether the mosaic segregated the array to its progeny. A predicted truncation of six amino acids of LET-23 (ref. 30) results from *let-23(sy1)*. *let-23(sy1)* is a stop mutation late in the *let-23* coding region³⁰.

Sequential signalling may be critical for the precise control of cell fate expression at the single-cell level; the anchor-cell signal may cause one cell to adopt the primary cell fate and the lateral signal may subsequently cause one cell (one each side of the primary cell) to adopt the secondary cell fate.

Previous experiments have shown that a Pn.p cell can sometimes express the secondary cell fate in the absence of an adjacent primary cell^{4,5}, suggesting that the anchor-cell signal can sometimes directly induce a Pn.p cell to express the secondary cell fate. Thus the lateral signal from the primary cell and medium levels of anchor-cell signal may be two ways to induce the expression of the secondary cell fate (Fig. 1c, right). The combined action of these two signals may ensure that P5.p and P7.p always express the secondary cell fate and thus may also ensure that the pattern of vulval cell fates is invariant in wild-type animals.

In addition to induction of the *C. elegans* vulva, cell signalling controls pattern formation in other developmental systems, such as induction of *Xenopus* mesoderm, formation of the vertebrate limb, patterning of *Drosophila* wing and segments, and establish-

ment of the *Drosophila* dorsoventral axis^{6,17,21}. To understand the mechanism(s) underlying pattern formation, it will be important to test further whether the inductive signals in each case function as morphogens or as the first inducers in a signalling cascade. □

Received 1 December 1994; accepted 28 March 1995.

- Horvitz, H. R. & Sternberg, P. W. *Nature* **351**, 535–541 (1991).
- Eisenmann, D. M. & Kim, S. K. *Curr. Opin. Genet. Dev.* **4**, 508–516 (1994).
- Kimble, J. *Devl Biol.* **87**, 286–300 (1981).
- Sternberg, P. W. & Horvitz, H. R. *Cell* **44**, 761–772 (1986).
- Thomas, J. H., Stern, M. J. & Horvitz, H. R. *Cell* **62**, 1041–1052 (1990).
- Hill, R. J. & Sternberg, P. W. *Nature* **358**, 470–476 (1992).
- Aroian, R. V., Koga, M., Mendel, J. E., Ohshima, Y. & Sternberg, P. W. *Nature* **348**, 693–699 (1990).
- Ferguson, E. L. & Horvitz, H. R. *Genetics* **110**, 17–72 (1985).
- Sulston, J. E. & Horvitz, H. R. *Devl Biol.* **82**, 41–55 (1981).
- Ferguson, E. L., Sternberg, P. W. & Horvitz, H. R. *Nature* **326**, 259–267 (1987).
- Herman, R. K. *Genetics* **108**, 165–180 (1984).
- Herman, R. K. *J. Neurogenet.* **5**, 1–24 (1989).
- Lackner, M. R., Kornfeld, K., Miller, L. M., Horvitz, H. R. & Kim, S. K. *Genes Dev.* **8**, 160–173 (1994).
- Aroian, R. V. & Sternberg, P. W. *Genetics* **128**, 251–267 (1991).
- Sternberg, P. W. *Nature* **335**, 551–554 (1988).

16. Greenwald, I. & Rubin, G. M. *Cell* **68**, 271–281 (1992).
17. Slack, J. M. W. *Curr. Biol.* **4**, 116–126 (1994).
18. Johnson, R. L., Riddle, R. D. & Tabin, C. J. *Curr. Opin. Genet. Dev.* **4**, 535–542 (1994).
19. Basler, K. & Struhl, G. *Nature* **368**, 208–214 (1994).
20. Ingham, P. W. *Curr. Biol.* **4**, 347–350 (1994).
21. Ferguson, E. L. & Anderson, K. V. *Cell* **71**, 451–461 (1992).
22. Sulston, J. E. & Horvitz, H. R. *Dev. Biol.* **56**, 110–156 (1977).
23. Sternberg, P. W. & Horvitz, H. R. *Cell* **58**, 679–693 (1989).
24. Miller, D. M., Niemeyer, C. J. & Chitkara, P. *Genetics* **135**, 741–753 (1993).
25. Sulston, J. E., Schierenberg, E., White, J. G. & Thomson, J. N. *Dev. Biol.* **100**, 64–119 (1983).
26. Fire, A. *EMBO J.* **5**, 2673–2680 (1986).
27. Mello, C. C., Kramer, J. M., Stinchcomb, D. & Ambros, V. *EMBO J.* **10**, 3959–3970 (1991).
28. Yuan, J. Y. & Horvitz, H. R. *Dev. Biol.* **138**, 33–41 (1990).
29. Austin, J. & Kimble, J. *Cell* **51**, 589–599 (1987).
30. Aroian, R. V., Lesa, G. M. & Sternberg, P. W. *EMBO J.* **13**, 360–366 (1994).

ACKNOWLEDGEMENTS. We thank D. Eisenmann, R. Hoskins, M. Lackner, J. Sisson and A. Villeneuve for their comments; members of S.K.K.'s laboratory for fruitful discussions; L. Miller for helpful discussions regarding methods; and P. Sternberg, D. Waring and D. Miller for providing us with the *let-23(sy1)* and *let-23(sy15)* strains, and the *pk7-13.8*, *C33C3* and *pNC4-21* clones. This work was supported by grants from the Lucille P. Markey charitable trust, the Searle Scholars Program and the NIH.

Male–female differences in fertility and blood pressure in ACE-deficient mice

John H. Krege^{*†}, Simon W. M. John[†],
Laura L. Langenbach[†], Jeffrey B. Hodgin[†],
John R. Hagaman[†], Eric S. Bachman[†],
J. Charles Jennette[†], Deborah A. O'Brien[‡]
& Oliver Smithies[†]

Departments of ^{*} Internal Medicine, [†] Pathology, [‡] Pediatrics and Cell Biology and Anatomy, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7525, USA

ANGIOTENSIN-CONVERTING enzyme (ACE) is a dipeptidyl carboxypeptidase that generates the vasoconstricting peptide angiotensin II and inactivates the vasodilating peptide bradykinin¹. The gene encoding ACE is composed of two homologous regions and codes for both a somatic and testis isoenzyme^{2–4}. Experiments with hypertensive rats^{5,6} and some^{7–9}, but not other^{10–13}, studies of humans suggest that sequences at or linked to the gene influence blood pressure. The testis-specific form of ACE has its own promoter within intron 12 (ref. 14), is encoded by the 3' region of the gene, and is found only in postmeiotic spermatogenic cells and sperm¹⁵. Its function is unknown¹⁶. Here we investigate the role of the *Ace* gene in blood pressure control and reproduction using mice generated to carry an insertional mutation that is designed to inactivate both forms of ACE. All homozygous female mutants were found to be fertile, but the fertility of homozygous male mutants was greatly reduced. Heterozygous males but not females had blood pressures that were 15–20 mm Hg less than normal, although both male and female heterozygotes had reduced serum ACE activity.

We used gene targeting to disrupt exon 14 of the *Ace* gene in a subline, BK4, of E14TG2a¹⁷ strain 129 embryonic stem cells (Fig. 1). Exon 14 encodes amino acids critical for the normal function of both testis and somatic ACE¹⁸. Male chimaeras derived from the stem cells were mated with C57BL/6J females to obtain heterozygous F₁ hybrids.

When the F₁ heterozygotes were mated to generate F₂ offspring of the three genotypes (+/+, +/- and -/-), litter sizes were normal (Table 1) but there was a marked deficiency of +/- and -/- weanlings. The proportions of +/+, +/- and -/- offspring (90+/- (38 m, 52 f), 124 +/- (58 m, 66 f), and 27 -/- (11 m, 16 f), where m is male and f is female) differed from the expected 1:2:1 mendelian proportions ($P < 0.001$ by χ^2) and the proportions of +/+ to +/- differed from 1:2

($P < 0.01$). The cause of these deviations from the expected values has not been determined.

The presence of an isoenzyme of ACE (testis ACE) in post-meiotic spermatogenic cells and sperm, suggests that testis ACE might be important for overall sperm function. In support of this possibility, the presence of an ACE inhibitor in the medium reduces the percentage of zona-free hamster oocytes penetrated by human sperm *in vitro*¹⁹. We therefore investigated the fertility of -/- females and -/- males (Table 1). Five of five -/- females were fertile, whereas only one of five -/- males sired

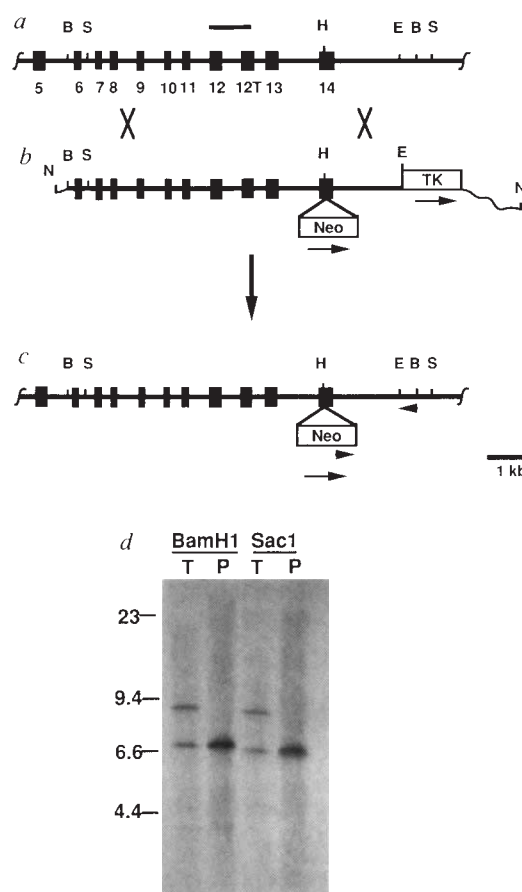


FIG. 1 Disruption of the *Ace* gene. *a*, The endogenous *Ace* locus from exon 5 to 14. Bar indicates the cloning probe; black boxes indicate exons; 12T denotes the testis-specific exon; B, E, H, N, and S indicate *Bam*HI, *Eco*RI, *Hind*III, *Not*I and *Sac*I sites. *b*, Targeting construct. Neo and TK indicate neomycin-resistance and thymidine kinase genes, oriented as indicated by the arrows. Wavy lines indicate plasmid sequences (not drawn to scale). *c*, The targeted *Ace* locus with the neomycin-resistance gene disrupting exon 14. Diagnostic PCR primers used to identify targeted cell lines provisionally are shown as left and right arrowheads. *d*, Southern blot showing expected band *Bam* HI = 8.5 kb and *Sac*I = 8.4 kb, for correctly targeted ES cell colony DNA (T) and *Bam* HI = 6.7 kb and *Sac*I = 6.6 kb, for parental cell line DNA (P). Size standards are shown (in kb) to the left.

METHODS. *Ace* DNA was cloned from strain-129 genomic DNA using the probe shown, made by PCR with primers designed from the mouse genomic sequence³. The targeting construct contains a ~5 kb *Bam*HI to *Hind*III fragment and a ~1.5 kb *Hind*III to *Eco*RI fragment inserted into the pPNT vector²⁴. The construct was linearized with *Not*I and targeted cells were produced as described²⁵. Colonies were provisionally identified by PCR using the primers shown in *c*. Southern blots of *Bam*HI- and *Sac*I-digested colony DNA were hybridized with the probe shown to confirm correct targeting. The targeted embryonic stem cells were injected into blastocysts to generate chimaeras. Four of the chimaeras mated to C57BL/6J females transmitted the mutant *Ace* gene to their agouti F₁ offspring. F₁ mice (+/-) having the *Ace* mutation were bred to obtain +/+, +/- and -/- F₂ mice.

TABLE 1 Breeding experiments with Ace mutant mice

Experiment	Mate	Litters	Pups
+/- Females (n=14) and males (n=7)	+/-	34	241*
-/- Females (n=5)†	+/+	2	8
	+/+	1	10
	+/-	2	10
	+/-	2	13
	+/-	2	13
		9	54
-/- Males (n=5)†	+/+	0	0
	+/+	1	2
	+/+	0	0
	+/+	0	0
	+/+	0	0
		1	2

* Litter size (7.1 pups per litter) was the same as normal. For comparison, offspring from mating heterozygotes for a duplicated angiotensinogen gene²⁰ were obtained in 1:2:1 proportions and litter size was 6.6 pups per litter.

† Homozygous mutant mice were with +/+ mates for 6 weeks, apart from one male that was with two +/+ females for 4 weeks. Two +/+ female mates that produced no litters with -/- males were subsequently tested with +/+ males, and both were found to be fertile.

a litter (two pups), demonstrating that the fertility of -/- males was substantially reduced.

The potency of -/- males was evaluated to assess the possibility that they were not mating. A single superovulated CD1 female was put with an individual -/- male and the vagina of the female was checked for a copulatory plug the following day. This test showed that five of five -/- males copulated successfully. Also, testis histology, sperm counts and sperm morphology of two -/- and two +/- males were not distinguishable from +/+. Our findings that -/- males have reduced fertility, despite being potent and having sperm of normal appearance, suggests that sperm produced by -/- males have reduced ability to fertilize ova.

TABLE 2 Serum ACE activities and blood pressures in Ace mutant mice

Study	Group	Females (n)	Males (n)	Combined (n)
Serum ACE activities (U/L) of F ₂ mice	+/+	377 ± 23 (n=11)	409 ± 37 (n=10)	392 ± 21 (n=21)
	+/-	248 ± 28* (n=10)	313 ± 22† (n=19)	291 ± 18* (n=29)
	-/-	40 ± 15‡ (n=4)	-39 ± 43‡ (n=7)	-10 ± 29‡ (n=11)
Tail-cuff pressures (mm Hg) of F ₂ mice	+/+	115 ± 1 (n=21)	119 ± 2 (n=13)	116 ± 1 (n=34)
	+/-	114 ± 2 (n=13)	104 ± 5* (n=7)	111 ± 2† (n=20)
	-/-	81 ± 8‡ (n=4)	85 ± 11§ (n=4)	83 ± 6‡ (n=8)
Mean arterial pressures (mm Hg) of F ₁ mice	+/+	132 ± 5 (n=7)	142 ± 5 (n=6)	137 ± 4 (n=13)
	+/-	134 ± 2 (n=7)	122 ± 6† (n=6)	129 ± 3 (n=13)

All experiments were performed in blinded fashion on mice fed Agway R-M-H3500 autoclaved feed. Results are shown as means ± s.e.m. and statistics for compared groups were by two-tailed *t*-test. ACE activities on serum (5 µl) obtained from retro-orbital blood diluted in 15 µl phosphate-buffered saline were by cleavage of the chromogenic tripeptide furylacryloylphenylalanyl-glycylglycine²² (Sigma Diagnostics) using the manufacturer's protocol with volumes reduced 5-fold. Significant differences were observed between genotypes but not between males and females of the same genotype. Tail-cuff blood pressures of F₂ mice aged 92 to 180 days were with a computerized, automated system (Visitech Systems) as described²³. The tail-cuff pressures of +/- males but not females were significantly lower than those of +/+. Mean arterial pressures were determined in F₁ mice aged 141 to 172 days as described²³. Catheters were inserted into the left carotid artery under general anaesthesia (2 mg ketamine and 0.2 mg xylazine per 25–30 g body weight, i.m.). Mice were allowed to recover for at least 5 h before measurements were taken, when they were resting and unrestrained. The mean arterial pressures of +/- males but not females were significantly lower than those of +/+ mice.

* The +/- group differed from the +/+ group (*P* < 0.01).

† The +/- group differed from the +/+ group (*P* < 0.03).

‡ The -/- group differed from the +/- group (*P* < 0.001).

§ The -/- group tended to differ from the +/- group (*P* = 0.10) and was significantly different from the +/+ group (*P* < 0.001).

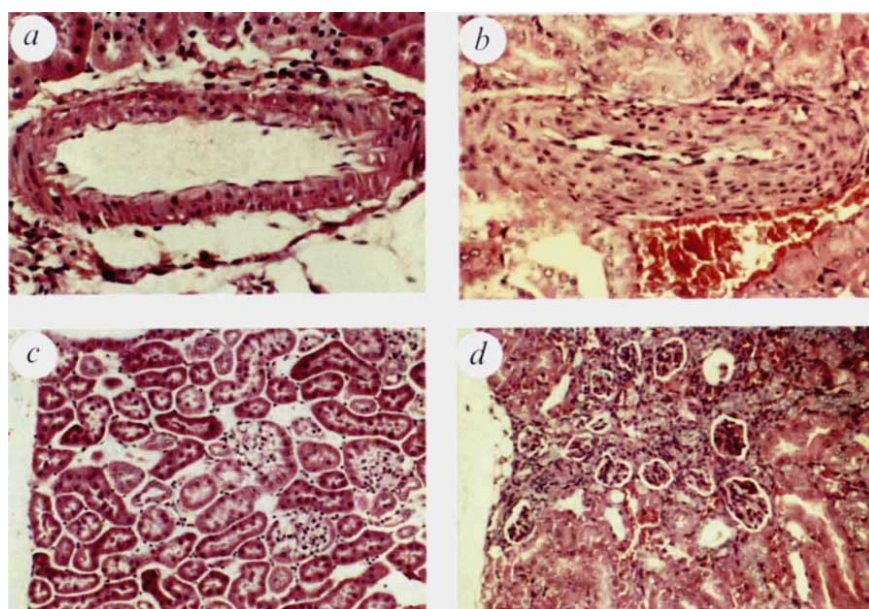


FIG. 2 Representative histology of +/+ and -/- mouse kidneys stained with haematoxylin and eosin (the kidneys from +/- mice were indistinguishable from +/+). a, Intrarenal artery from a +/+ mouse demonstrating the normal thickness of the muscularis (35 mm; magnification, ×73). b, Comparable intrarenal artery from a -/- mouse showing marked mural hypercellularity and thickening with narrowing of the lumen (35 mm; magnification, ×73). c, Normal subcapsular renal cortex from a +/+ mouse (35 mm; magnification, ×41). d, Subcapsular wedge-shaped zone of cortical atrophy with shrinkage of tubules, chronic inflammation, and crowding of glomeruli in a -/- mouse, possibly secondary to ischaemia caused by the narrowing of arterial lumens as in b (35 mm; magnification, ×41).

Adult mice of all genotypes appeared to be outwardly healthy, although adult $-/-$ mice (3 males, 2 females) had abnormal kidneys with cortical thinning, focal areas of atrophy, and renal vascular changes consisting of thickened walls composed of an increased number of disorganized cells (Fig. 2). These changes are identical to those observed in mice lacking a functional angiotensinogen gene²⁰. Other tissues, including the heart, lungs and jejunum, appeared to be histologically normal.

Serum ACE activities in $+/+$, $+/-$ and $-/-$ mice were determined (Table 2). Heterozygous males and females had serum activities 23 and 35% lower than $+/+$, whereas homozygous mutant animals had ACE activities not distinguishable from zero.

We determined the tail-cuff blood pressures of F_2 mice and mean arterial blood pressures of F_1 mice (Table 2). Heterozygosity for the *Ace* gene disruption lowers the blood pressure of males by 15–20 mm Hg; heterozygous females have normal blood pressure. Thus, although both male and female $+/-$ mice have reduced serum ACE levels, only the males show a decrease in blood pressure. The blood pressure of male and female $-/-$ mice is reduced by ~ 35 mm Hg.

We do not know why the blood pressure of male and female mice should be differently affected by heterozygosity for disruption of the *Ace* gene. The mutation shown in Fig. 1 was designed to inactivate both the somatic and testis forms of ACE, so the male–female difference may stem from a function mediated by testis ACE. Mutations designed to modify the expression of testis and somatic ACE separately should help to clarify the individual roles of the two forms. Another consideration is that some strains of mice, including strain 129 but not C57BL/6J, have a renin gene that is androgen-sensitive in the submandibular gland²¹.

Linkage studies in rodents^{5,6} and case-control studies in humans^{7–9} have suggested that sequences at, or linked to, the ACE gene can influence blood pressure. Our gene targeting approach to studying the regulation of blood pressure has the advantage that all F_1 mice are genetically identical apart from their *Ace* genes and sex chromosomes, so the differences in blood pressure in male $+/+$ and $+/-$ F_1 mice are directly caused by the *Ace* mutation and are not due to any other linked or unlinked genetic effects.

The finding that male mice homozygous for a disrupted *Ace* gene have impaired fertility may have important implications for human fertility and contraception. Also, as heterozygosity for the disrupted *Ace* gene lowers the blood pressure of males but not of females, the effects of gender and the human *Ace* gene in blood pressure regulation warrant further investigation. $\square$

Received 16 February; accepted 27 March 1995.

1. Erdős, E. G. *Hypertension* **16**, 363–370 (1990).
2. Bernstein, K. E., Martin, B. M., Edwards, A. S. & Bernstein, E. A. *J. biol. Chem.* **264**, 11945–11951 (1989).
3. Soubrier, F. et al. *Proc. natn. Acad. Sci. U.S.A.* **85**, 9386–9390 (1988).
4. Sibony, M., Gasc, J.-M., Soubrier, F., Alhenc-Gelas, F. & Corvol, P. *Hypertension* **21**, 827–835 (1993).
5. Jacob, H. J. et al. *Cell* **67**, 213–224 (1991).
6. Hilbert, P. et al. *Nature* **353**, 521–529 (1991).
7. Duru, K., Farrow, S., Wang, J.-M., Lockette, W. & Kurtz, T. Am. *J. Hypertens.* **7**, 759–762 (1994).
8. Morise, T., Takeuchi, Y. & Takeda, R. *Lancet* **343**, 125 (1994).
9. Zee, R. Y. L., Lou, Y.-K., Griffiths, L. R. & Morris, B. J. *Biochem. biophys. Res. Commun.* **184**, 9–15 (1992).
10. Harrop, S. B. et al. *Hypertension* **21**, 455–460 (1993).
11. Higashimori, K. et al. *Biochem. biophys. Res. Commun.* **191**, 399–404 (1993).
12. Schmidt, S., van Hooft, I. M. S., Grobbee, D. E., Ganten, D. & Ritz, E. *J. Hypertens.* **11**, 345–348 (1993).
13. Jeunemaitre, X., Lifton, R. P., Hunt, S. C., Williams, R. R. & Lalouel, J.-M. *Nature Genet.* **1**, 72–75 (1992).
14. Howard, T. E., Shai, S. Y., Langford, K. G., Martin, B. M. & Bernstein, K. E. *Molec. cell. Biol.* **10**, 4294–4302 (1990).
15. Langford, K. G., Zhou, Y., Russell, L. D., Wilcox, J. N. & Bernstein, K. E. *Biol. Reprod.* **48**, 1210–1218 (1993).
16. Ehlers, M. R. W. & Riordan, J. F. *Biochemistry* **28**, 5313–5318 (1989).
17. Hooper, M., Hardy, K., Handyside, A., Hunter, S. & Monk, M. *Nature* **326**, 292–295 (1987).
18. Sen, I., Kasturi, S., Abdul Jabbar, M. & Sen, G. C. *J. biol. Chem.* **268**, 25748–25754 (1993).
19. Foresta, C., Mioni, R., Rossato, M., Varotto, A. & Zorzi, M. *Int. J. Androl.* **14**, 333–339 (1991).
20. Kim, H. S. et al. *Proc. natn. Acad. Sci. U.S.A.* (in the press).

21. Wilson, C. M., Cherry, M., Taylor, B. A. & Wilson, J. D. *Biochem. Genet.* **19**, 509–523 (1981).
22. Holmquist, B., Bünning, P. & Riordan, J. F. *Analyt. Biochem.* **95**, 540–548 (1979).
23. Kregel, J. H., Hodgins, J. B., Hagaman, J. R. & Smithies, O. *Hypertension* (in the press).
24. Tybulewicz, V. L. J., Crawford, C. E., Jackson, P. K., Bronson, R. T. & Mulligan, R. C. *Cell* **65**, 1153–1163 (1991).
25. Smithies, O. & Kim, H. S. *Proc. natn. Acad. Sci. U.S.A.* **91**, 3612–3615 (1994).

ACKNOWLEDGEMENTS. We thank M. Eddy, F. Fiedorek, F. French, S. Hiller, H.-S. Kim, K. Kluckman, D. Lee, J. Lynch, N. Maeda, S. Morham, J. Moyer and C. Sapienza for their help. These studies were supported by grants from the NIH. J.H.K. is a Howard Hughes Medical Institute Physician Research Fellow.

Serial triggering of many T-cell receptors by a few peptide–MHC complexes

Salvatore Valitutti, Sabina Müller, Marina Cella, Elisabetta Padovan & Antonio Lanzavecchia

Basel Institute for Immunology, Grenzacherstrasse 487, CH 4005, Basel, Switzerland

T LYMPHOCYTES can recognize and be activated by a very small number of complexes of peptide with major histocompatibility complex (MHC) molecules displayed on the surface of antigen-presenting cells (APCs)^{1,2}. The interaction between the T-cell receptor (TCR) and its ligand has low affinity and high off-rate^{3–6}. Both findings suggest that an extremely small number of TCRs must be engaged in interaction with APCs and raise the question of how so few receptors can transduce an activation signal. Here we show that a small number of peptide–MHC complexes can achieve a high TCR occupancy, because a single complex can serially engage and trigger up to ~ 200 TCRs. Furthermore, TCR occupancy is proportional to the T cell's biological response. Our findings suggest that the low affinity of the TCR can be instrumental in enabling a small number of antigenic complexes to be detected.

An unresolved question in T-cell activation is the stoichiometry and kinetics of the TCR–ligand interaction *in vivo*. Although the number of peptide–MHC complexes on an APC can be directly measured by biochemical methods^{1,2}, the number of TCRs engaged at the level of specific T–APC conjugates remains to be established.

We have investigated the possibility of using TCR/CD3 downregulation to estimate TCR occupancy *in vivo*. It has been shown that in T cells TCR/CD3 complexes are internalized and recycled constitutively^{7,8} but, when triggered by crosslinked anti-CD3 antibodies, they are downregulated following serine phosphorylation of the CD3- γ chain, which targets them for lysosomal degradation^{9–11}.

To determine the extent of TCR downregulation induced by specific antigen, we formed conjugates between T-cell clones and autologous Epstein–Barr virus (EBV)-transformed B cells pulsed with different concentrations of the antigenic peptide and measured the level of TCR/CD3 complexes after different times by staining with anti-CD3 antibody. Figure 1 shows that T cells conjugated with peptide-pulsed APCs undergo a profound and time-dependent downregulation of the TCR/CD3 complex. Incubation with APCs pulsed with a high peptide concentration (10 μ M) results in downregulation of most surface TCRs ($>90\%$). Strikingly, even APCs pulsed with a low peptide concentration (25 nM, which is limiting for T-cell activation) downregulate a considerable proportion of TCRs (up to 50%). The time course of TCR downregulation is similar at high and low peptide concentrations, reaching a plateau after ~ 2 hours (Fig. 1b).

To investigate whether downregulation involves only those TCRs that have been specifically triggered or a random fraction of all TCRs, we used dual-receptor T-cell clones expressing two different $\alpha\beta$ dimers which can be distinguished using specific

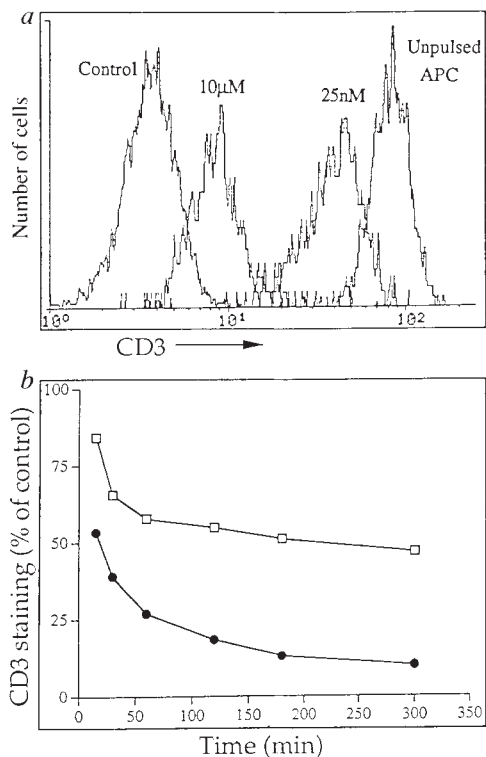


FIG. 1 T cells conjugated with peptide-pulsed APCs undergo an antigen-dependent downregulation of the TCR/CD3 complex. *a*, CD3 staining of a T-cell clone that had been conjugated for 5 h with autologous EBV-B cells either unpulsed or pulsed with 25 nM or 10 μM peptide. In a parallel experiment the same peptide-pulsed APCs induced 32% and 100% of the maximum IFN- γ production respectively. *b*, Time course of CD3 downregulation in the same T-cell clone conjugated with autologous EBV-B cells pulsed with 25 nM (□) or 20 μM (●) peptide. 100% value corresponds to the median CD3 fluorescence of the clone conjugated with unpulsed EBV-B cells. This value did not change with time of incubation. The level of CD3 did not recover for at least 24 hours after conjugation.

METHODS. A DRBI*1104-restricted T-cell clone (KS 140) specific for tetanus toxin (TT) peptide 830–843 (QYIKANSKFIGITE) was used¹⁸. Autologous EBV-B cells were pulsed for 2 h at 37 °C with different concentrations of TT830–843 in RPMI-5% FCS. During the last 10 minutes 1 μM BCECF-AM (2',7-bis-(carboxyethyl)-5(6')-carboxyfluorescein) (Calbiochem, San Diego, CA), was added and the cells were washed 4 times. T cells were mixed with EBV-B cells at a 1:2 ratio in 200 μl RPMI-5% FCS in U-bottom microplates, centrifuged to allow conjugate formation and incubated at 37 °C. At different times the cells were resuspended, washed in PBS containing 0.5 mM EDTA to break the conjugates and stained with anti-CD3 (OKT3, ATCC, Rockville, MD) followed by a phycoerythrin (PE)-labelled goat anti mouse Ig (SBA, Birmingham, AL). The CD3 fluorescence was analysed on a FACScan (Becton-Dickinson, Mountain View, CA). EBV-B cells were gated out using both FSC/SSC parameters and green BCECF fluorescence. Dead cells were excluded by propidium iodide staining. IFN- γ was measured in the culture supernatant using an ELISA assay as described (ref. 19). Data are from 1 representative experiment out of 5.

antibodies¹². When a V β 2⁺V β 8⁺ clone is triggered by the V β 2-specific ligand TSST, only the V β 2⁺ TCRs are downregulated, whereas the V β 8⁺ TCRs remain unaffected (Fig. 2*a*). In a V β 2⁺V β 3⁺ alloreactive T-cell clone, a selective downregulation of V β 3⁺ TCRs is induced by the specific alloantigen; triggering with plastic-bound anti-V β 2 antibody results in complete downregulation of V β 2⁺ TCRs without affecting the surface expression of V β 3⁺ TCRs (Fig. 2*b*). In a T-cell clone expressing V α 2 and V α 24, stimulation by plastic-bound anti-V α antibodies results in downregulation of the cognate V α , whereas the other V α is affected to a much lesser extent (Fig. 2*c*). Taken together, these results indicate that downregulation involves primarily the TCRs that are specifically triggered, with the bystander TCRs being essentially unaffected unless a very strong stimulus is applied, and then only slightly. As we consider it to be unlikely that an antigen-independent homotypic clustering of TCR/CD3 complexes may be responsible for these results, we conclude that TCR downregulation can be used to measure TCR occupancy *in vivo*.

To estimate the stoichiometry of peptide-MHC/TCR interaction in specific conjugates, we measured at different antigen concentrations the number of complexes per APC and the number of TCRs downregulated after T-APC interaction. In one series of experiments, EBV-B cells were pulsed with different concentrations of ¹²⁵I-labelled peptide and the number of peptide-DR complexes per cell was calculated at each peptide concentration (Fig. 3*a*). In parallel experiments, the fraction of TCRs downregulated by APCs pulsed in the same conditions was measured (Fig. 3*b*). From comparison of the two curves it is estimated that APCs pulsed with high peptide concentrations (20 μM) display ~7,500 complexes and induce downregulation of 93% of TCRs. Strikingly, APCs pulsed with a low peptide concentration (50 nM) display only ~100 peptide-DR complexes, yet they downregulate 62% of TCRs. The relationship between the number of peptide-DR complexes per APC and the number of TCRs downregulated per T cell is shown in Fig. 3*c*. This plot clearly shows that each peptide-DR complex must engage a large number of TCRs in successive rounds. This effect is dramatic at low complex density, where ~100 complexes can trigger up to 18,000 TCRs, but is less marked at high complex

density, indicating that a single peptide-DR must be able to trigger 180 TCRs in successive rounds. This figure may increase at lower complex density and could be an underestimate as it is unlikely that all complexes present on an APC may be available to the responding T cell.

We next asked whether TCR occupancy, as measured by TCR downregulation, would reflect the extent of the T cell's biological response. Figure 4*a* shows that the production of interferon (IFN)- γ by T cells correlates with the number of TCRs downregulated. This correlation was also studied in T cells stimulated by partial agonists, namely substituted peptides that bind to DR molecules with the same affinity but have less capacity to stimulate T cells¹³. Figure 4*b* shows that, although the partial agonist is ~1,000-fold less effective than the antigenic peptide, a similar relationship between TCR downregulation and biological response exists. Taken together, these results indicate that (1) the extent of the T-cell response reflects TCR occupancy, and (2) substituted peptides may behave as partial agonists because they trigger fewer TCRs.

The high-sensitivity-low-affinity paradox of T-cell antigen recognition^{14,15} can be explained by our finding that a small number of peptide-MHC complexes can serially trigger a much larger number of TCRs, leading to an amplified and sustained signal. The high off-rate of TCR can be instrumental to this mechanism, because it allows a single peptide-MHC complex to engage many TCRs in successive rounds of ligation, triggering and dissociation. Our estimate of peptide-MHC re-usage is consistent with the estimated half-lives of murine TCR-ligand complexes^{5,14}.

Although we cannot formally rule out the possibility that TCR crosslinking might occur in T-APC conjugates (especially at high epitope density), our results suggest that under physiological conditions T cells are triggered by sequential monovalent engagement of many TCRs by a small number of ligands. According to this model, a minimum time of ligation is required to assemble the multicomponent complex that transduces an activatory signal and is eventually downregulated.

Our model offers an explanation for the phenomenon of TCR antagonism. It is possible that antagonistic complexes may rapidly engage and incompletely trigger a large number of TCRs

FIG. 2 TCR downregulation in dual-receptor T cells is selective for the TCRs that are specifically triggered. **a**, A $V\beta 2^+ V\beta 8^+$ T-cell clone (2.8.12) was conjugated with TSST-pulsed or untreated EBV-B cells and stained with anti- $V\beta 2$ (left) or anti- $V\beta 8$ (right). **b**, An alloreactive $V\beta 2^+ V\beta 3^+$ T-cell clone (2.3.33) was conjugated with specific or irrelevant EBV-B stimulator cells or cultured in wells coated with anti- $V\beta 2$ antibody. The cells were stained with anti- $V\beta 3$ (left) or anti- $V\beta 2$ (right). **c**, A $V\alpha 2^+ V\alpha 24^+$ T-cell clone (24.2.7) was cultured in tissue-culture wells coated with either anti- $V\alpha 2$ or anti- $V\alpha 24$ antibody and stained with anti- $V\alpha 2$ (left) or anti- $V\alpha 24$ (right).

METHODS. The following mouse monoclonal antibodies were used: anti- $V\beta 2$ (MPB2/D5, IgG1; ref. 20), anti- $V\beta 3$ (LE-89, IgG2a), anti- $V\beta 8$ (56C5, IgG2a) (Immunotech, Marseille), anti- $V\alpha 2$ (IgG2a; T-Cell Diagnostic, Cambridge, Massachusetts) and anti- $V\alpha 24$ (C15, IgG1; ref. 12). Dual-receptor T-cell clones carrying two $V\alpha$ or two $V\beta$ have been described^{12,21}. Clone 2.8.12 was conjugated with autologous EBV-B cells either unpulsed or pulsed for 1 h with 10 ng ml^{-1} TSST-1 (Toxin Technology). Clone 2.3.33 was conjugated with stimulatory or control EBV-B cells or cultured in wells coated with anti- $V\beta 2$ antibody (1:300 ascitis in PBS). Clone 24.2.7 was cultured in tissue-culture wells coated with anti- $V\alpha 2$ or anti- $V\alpha 24$ antibodies ($2 \mu\text{g ml}^{-1}$ in PBS). After 5 h the cells were collected and stained with a saturating concentration of the relevant anti- $V\alpha$ or - $V\beta$ antibodies followed by isotype-specific second-step reagents. Data are from 1 representative experiment out of at least 3 for each dual-receptor T-cell clone.

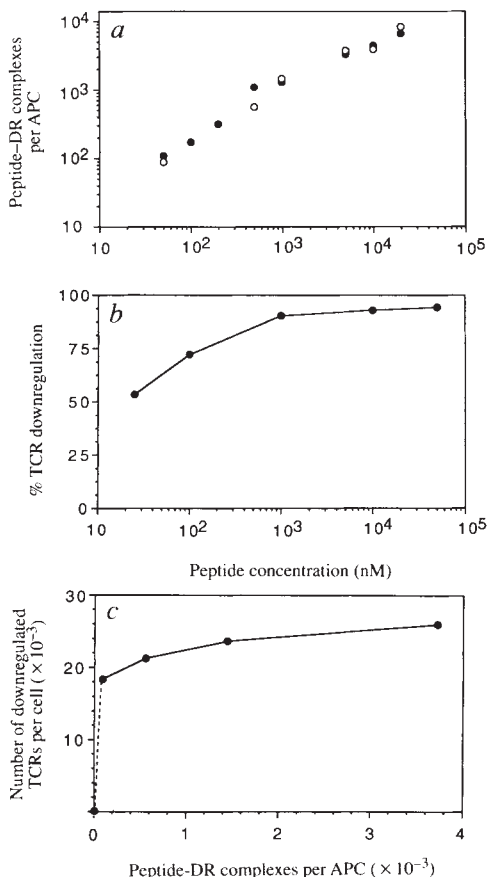
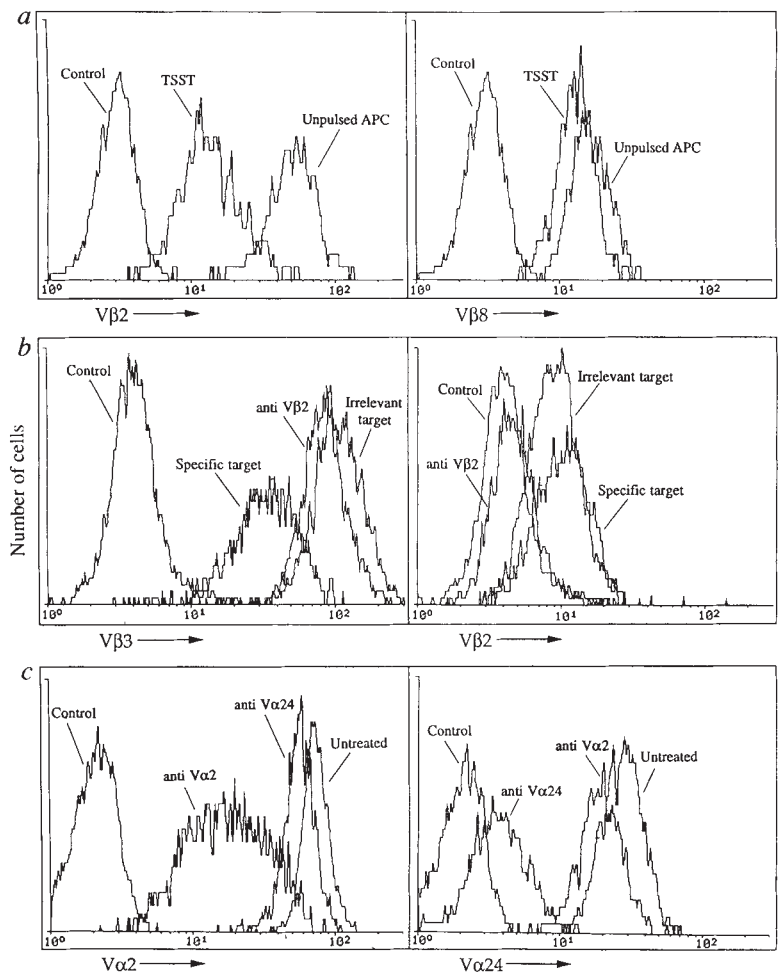


FIG. 3 A small number of peptide-MHC complexes can engage and downregulate a large number of TCRs. **a**, Number of peptides bound to DR molecules in EBV-B cells pulsed for 2 h at 37°C with different peptide concentrations. Data are from two experiments ($\bullet$, $\circ$). **b**, TCR downregulation induced by EBV-B cells pulsed for 2 h at 37°C with different peptide concentrations. **c**, Relationship between the number of peptide-DR complexes on APC and the number of downregulated TCRs per T cell. The experimental points correspond to APC pulsed with 50, 500, 1,000 and 5,000 nM peptide. Quantitation of peptide-DR complexes and TCR downregulation were done as in **a** and **b**.

METHODS. Quantitation of peptide-DR complexes: TT830-843 was labelled with ^{125}I using the chloramine-T method. 8×10^6 DRB1*1104 EBV-B cells were pulsed with different concentrations of ^{125}I -labelled peptide for 2 h at 37°C in culture medium in the same conditions used for T-cell-stimulation experiments, washed and lysed as described²². Three consecutive rounds of immunoprecipitation were done using anti-DR antibody L243 (ATCC) and the radioactivity in the precipitates counted. An irrelevant antibody of the same isotype (OKT3, ATCC) precipitated less than 10% of the radioactivity precipitated by L243. When the precipitates were eluted in sample buffer at room temperature²³ and run on a 15% SDS-PAGE, the radioactivity was associated with compact DR heterodimers. The number of peptides per APC was calculated from the total counts in the precipitates, the number of cells and the specific activity of the peptide (5×10^4 and 6.5×10^4 c.p.m. ng^{-1} in the two experiments shown). The results are from 2 experiments out of 5 using different EBV-B cells, with comparable results. TCR downregulation: TCR downregulation was measured 5 h after conjugation, as for Fig. 1. Data are from 1 experiment done in duplicate and representative of 7. The standard deviation at each point was less than 3%. Similar levels of CD3 downregulation were obtained using EBV-B cells that had been pulsed under the same conditions and fixed with glutaraldehyde before the assay. Quantitation of TCR/CD3 complexes: The number of TCR/CD3 complexes was measured by FACS analysis of T cells stained with saturating concentrations of FITC-labelled antibodies of known fluorescein/protein ratio and comparing the staining with a standard curve of microbeads labelled with defined numbers of fluorescein molecules (Flow Cytometry Standard Corp.) as described in ref. 24. Comparable estimates ($\sim 3 \times 10^4$) were obtained using anti-CD3 (TR66) and anti-TCR $\alpha\beta$ (BMA031; T-cell Diagnostics).

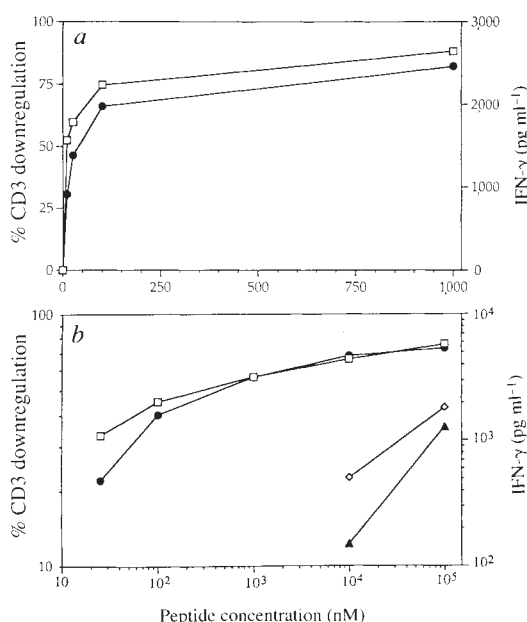


FIG. 4 IFN- γ production in T cells stimulated by antigenic peptide or a partial agonist reflects the extent of TCR downregulation. **a**, TCR downregulation ($\square$) and production of IFN- γ ($\bullet$) in clone KS 140 conjugated for 5 h with autologous EBV-B cells pulsed for 2 h at 37 °C with different concentrations of TT830–843. Data are from one representative experiment out of six using clone KS 140. A similar correlation between TCR downregulation and IFN- γ production was observed with 6 additional T-cell clones specific for different peptides. **b**, TCR downregulation ($\square$, $\diamond$) and production of IFN- γ ($\bullet$, $\blacktriangle$) by T-cell clone ALP24.3 (DRBI*1302-restricted; ref. 25) stimulated with autologous EBV-B cells pulsed for 2 h at 37 °C with different concentrations of either TT830–843 ($\square$, $\bullet$) or of a substituted peptide (I $\rightarrow$ Q 839) ($\diamond$, $\blacktriangle$) which binds with the same affinity to DRBI*1302 but has partial agonistic activity (O. Acuto, personal communication). Data are from 1 representative experiment out of 4.

and thus effectively compete with agonistic complexes, an idea supported by recent evidence that antagonists induce an altered pattern of TCR- ζ chain phosphorylation and fail to activate ZAP-70 (refs 16, 17).

The re-use of peptide–MHC ligands in order to achieve efficient T-cell activation implies that effective antigenic complexes must have an optimum affinity which allows efficient TCR triggering and complex recycling. An affinity less than that of an optimally stimulatory complex might lead to a smaller fraction of TCR being triggered and possibly to the simultaneous generation of 'spoiled' complexes, thus accounting for partial agonism and antagonism. Conversely, an increase in affinity could result in less stimulation, because the lower off-rate may prevent re-usage. This unique mechanism appears to have developed to ensure sensitive detection of the small number of antigenic determinants present on the surface of cells. $\square$

Received 19 December 1994; accepted 7 March 1995.

1. Harding, C. V. & Unanue, E. R. *Nature* **346**, 574–576 (1990).
2. Demotz, S., Grey, H. M. & Sette, A. *Science* **249**, 1028–1030 (1990).
3. Weber, S., Trautwein, A., Oliveri, F., Gerhard, W. & Karjalainen, K. *Nature* **356**, 793–796 (1992).
4. Matsui, K. et al. *Science* **254**, 1788–1791 (1991).
5. Corr, M. et al. *Science* **265**, 946–949 (1994).
6. Sykulev, Y. et al. *Immunity* **1**, 15–22 (1994).
7. Krangel, M. S. *J. exp. Med.* **165**, 1141–1159 (1987).
8. Minami, Y., Samelson, L. E. & Klausner, R. D. *J. biol. Chem.* **262**, 13342–13347 (1987).
9. Cantrell, D. A., Davies, A. A. & Crumpton, M. J. *Proc. natn. Acad. Sci. U.S.A.* **82**, 8158–8162 (1985).
10. Luton, F., Buferne, M., Davoust, J., Schmitt-Verhulst, A.-M. & Boyer, C. *J. Immun.* **153**, 63–72 (1994).
11. Dietrich, J., Hou, X., Wegener, A. M. & Geisler, C. *EMBO J.* **13**, 2156–2166 (1994).
12. Padovan, E. et al. *Science* **262**, 422–424 (1993).
13. Evavold, B. D., Sloan-Lancaster, J. & Allen, P. M. *Immun. Today* **14**, 602–609 (1993).
14. Williams, A. F. & Beyers, A. D. *Nature* **356**, 746–747 (1992).

15. Karjalainen, K. *Curr. Opin. Immun.* **6**, 9–12 (1994).
16. Sloan-Lancaster, J., Shaw, A. S., Rothbard, J. B. & Allen, P. M. *Cell* **79**, 913–922 (1994).
17. Madrenas, J. et al. *Science* **267**, 515–517 (1995).
18. Panina, P. et al. *Eur. J. Immun.* **19**, 2237–2242 (1989).
19. Valitutti, S. et al. *J. exp. Med.* **181**, 577–584 (1995).
20. Clarke, G. R., Reyburn, H., Lancaster, F. C. & Boylston, W. *Eur. J. Immun.* **24**, 837–842 (1994).
21. Padovan, E. et al. *J. exp. Med.* **181**, 1587–1591 (1995).
22. Lanzavecchia, A., Reid, P. A. & Watts, C. *Nature* **357**, 249–252 (1992).
23. Germain, R. N. & Hendrix, L. R. *Nature* **353**, 134–139 (1991).
24. Vogt, R. F., Cross, G. D., Henderson, L. O. & Phillips, D. L. *Cytometry* **10**, 294–302 (1989).
25. Boitel, B. et al. *J. exp. Med.* **175**, 765–777 (1992).

ACKNOWLEDGEMENTS. We thank O. Acuto, K. Karjalainen, E. Palmer and C. Watts for critical reading and comments, M. Dessing for help in flow cytometry, and A. Necker for anti-V β monoclonal antibodies. The Basel Institute for Immunology was founded and is supported by F. Hoffmann-La Roche, Basel, Switzerland.

The receptor DEC-205 expressed by dendritic cells and thymic epithelial cells is involved in antigen processing

Wanping Jiang*, William J. Swiggard†, Christine Heufler*, Michael Peng†, Asra Mirza†, Ralph M. Steinman† & Michel C. Nussenzweig*§||

* Laboratory of Molecular Immunology, § Howard Hughes Medical Institute, and † Laboratory of Cellular Physiology, The Rockefeller University, 1230 York Avenue, New York, New York 10021, USA

DENDRITIC cells and thymic epithelial cells perform important immunoregulatory functions by presenting antigens in the form of peptides bound to cell-surface major histocompatibility complex (MHC) molecules to T cells^{1–3}. Whereas B cells are known to present specific antigens efficiently through their surface immunoglobins, a comparable mechanism for the capture and efficient presentation of diverse antigens by dendritic cells and thymic epithelial cells has not previously been described. We show here that their antigen-presentation function is associated with the high-level expression of DEC-205, an integral membrane protein homologous to the macrophage mannose receptor and related receptors which are able to bind carbohydrates and mediate endocytosis. DEC-205 is rapidly taken up by means of coated pits and vesicles, and is delivered to a multivesicular endosomal compartment that resembles the MHC class II-containing vesicles implicated in antigen presentation. Rabbit antibodies that bind DEC-205 are presented to reactive T-cell hybridomas 100-fold more efficiently than rabbit antibodies that do not bind DEC-205. Thus DEC-205 is a novel endocytic receptor that can be used by dendritic cells and thymic epithelial cells to direct captured antigens from the extracellular space to a specialized antigen-processing compartment.

Dendritic cells are a unique class of leukocytes whose primary function is to capture antigens, process them and present them to T cells¹. Interactions between dendritic cells and specific T cells in the peripheral immune system lead to the induction of immune responses, whereas in the thymus presentation by dendritic cells leads to negative selection². Like dendritic cells, thymic epithelial cells present MHC-bound peptides to T cells, but instead of inducing T-cell activation or negative selection, thymic epithelial cells direct positive selection³. Thus a fundamental functional requirement for both dendritic cells and thymic epithelial cells is the uptake and processing of antigen, yet neither cell type is known to express receptors that are specialized for antigen capture or presentation.

† Present address: University of Innsbruck, Austria.

|| To whom correspondence should be addressed.

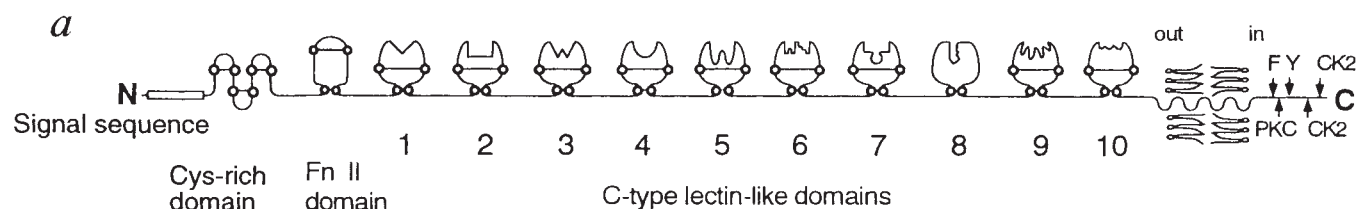


FIG. 1. Sequence of DEC-205. **a**, Schematic representation of DEC-205. **b**, The predicted amino-acid sequence of DEC-205 is aligned with the sequences of the bovine PLA₂ receptor (bv PLA2R) and the human macrophage mannose receptor (huMφmR). Amino-acid positions are shaded where there is identity among all three proteins. Protein domains are separated, and consensus amino acids that define C-type CRDs are indicated below the relevant sequence as follows: invariant amino acids are shown in single-letter code; θ , aliphatic; χ , aliphatic or aromatic; ϕ , aromatic; Z, E or Q; B, D or N; Ω , D or N or E or Q. The

two missing cysteines in CRD 8 are highlighted with an asterisk. Peptide sequences determined by automated Edman degradation of purified DEC-205 protein are overlined and numbered (N indicates amino terminal, T indicates peptides generated with trypsin, and L indicates peptides generated with endoproteinase lys-C). Symbols: N, amino terminal; Cys-rich domain, cysteine rich N-terminal domain; Fn II domain, fibronectin type II domain; PKC, potential site for protein kinase C mediated phosphorylation; CK2, potential site for casein kinase 2 mediated phosphorylation; Y, tyrosine; C, carboxyl terminus.

We attempted the molecular characterization of DEC-205, a cell-surface protein (M_r 205K) recognized by the NLDC-145 monoclonal antibody⁴, because it is expressed predominantly on dendritic cells and thymic epithelial cells (W. J. Swiggard, M.A., M.C.N. and R.M.S., submitted for publication). By using oligonucleotide probes based on the protein sequence, 12 complementary DNA clones were retrieved from 3 separate thymic and dendritic cell cDNA libraries. All of the cDNA clones were derived from the same messenger RNA. Clones containing the putative 5' end of the DEC-205 cDNA encoded the N-terminal peptide of the DEC-205 protein (W.J.S., M.A., M.C.N. and R.M.S., submitted), which was preceded by a hydrophobic leader consistent with a signal sequence, and the composite cDNA included all 29 unambiguous DEC-205 peptide sequences (Fig. 1). A 7.5-kb mRNA that corresponds to this cDNA was expressed at high levels in dendritic cells, thymus and lymph nodes, a pattern that corresponds to the histochemical staining pattern of the NLDC-145 monoclonal antibody⁴ (Fig. 2a). Finally, the cloned cDNA was shown to encode the antigen recognized by the NLDC-145 monoclonal antibody by western blotting of DEC-205-GST (glutathione *S*-transferase) fusion proteins expressed in bacteria (Fig. 2b).

When the sequence of DEC-205 was aligned with known proteins in the database, it was found to be homologous to the macrophage mannose receptor (MMR)^{5,7} and the phospholipase A₂ (PLA₂) receptors of rabbit skeletal muscle⁸ and bovine pancreas⁹ (Fig. 1). All known members of this family¹⁰ are type-I transmembrane proteins with short cytoplasmic domains that mediate adsorptive endocytosis^{6,9,11}. The ectodomain of MMR family proteins has a distinctive cysteine-rich N-terminal region, followed by a fibronectin type-II repeat, and eight Ca²⁺-dependent carbohydrate recognition domains (C-type CRDs)^{5,8,9}. DEC-205 diverges from this pattern only in that it has ten instead of the usual eight C-type CRDs (Fig. 1). The functions of the cysteine-rich domain and the fibronectin repeat have not been defined. In contrast, there is extensive experimental evidence that the C-type CRDs are carbohydrate-binding domains (reviewed in ref. 11), and both the MMR and rabbit PLA₂ receptors bind carbohydrates^{8,10}. The mannose contact residues defined for the C-type CRDs in rat serum mannose-binding protein A and E-selectin¹² are not conserved in DEC-205. However, additional mechanisms must exist for carbohydrate binding by C-type CRDs, because sequence features initially defined as 'essential' for carbohydrate binding are also absent from the CRDs of NKR-P1, Ly49A and rabbit PLA₂ receptors, all of which bind avidly to carbohydrate ligands^{8,13,14}. CRDs are found in over 50 known proteins¹⁰, but the CRDs in DEC-205 are most closely related to those found in the bovine and rabbit PLA₂ receptors, and the MMRs (34.6% identity with bovine PLA₂ receptor, and 26.7% identity with human MMR). At least two of the CRDs in the MMR are known to bind mannose¹¹,

and grouping of several CRDs increases the affinity of the MMR for carbohydrate ligands¹⁵. The same mechanism may enable DEC-205 to increase both the affinity and the diversity of carbohydrates bound by this receptor.

In the absence of information on the ligands bound by DEC-205, we investigated receptor function using a combination of monoclonal and polyclonal anti-DEC-205 antibodies. The conserved aromatic amino acids in the cytoplasmic domain of DEC-

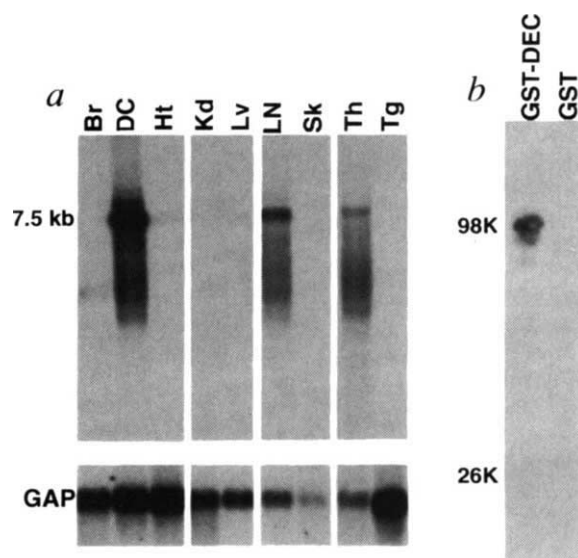


FIG. 2 Expression of DEC-205 in mouse tissues and *Escherichia coli*. **a**, Northern blot of poly(A)⁺ RNA extracted from the indicated tissues; the size of the RNA is indicated in kilobases. **b**, Western blot of *E. coli* extracts from cells transformed with a plasmid encoding either a GST-DEC-205 fusion protein or GST alone; M_r values are indicated. For northern blots, 2 μ g mRNA was electrophoresed in 0.8% agarose formaldehyde gels. Samples were transferred to nylon membranes and probed with an antisense RNA probe that spanned nucleotide positions 3688–5200 in the DEC-205 cDNA. The blot was subsequently stripped and rehybridized with a glyceraldehyde-3-phosphate dehydrogenase probe (GAP) as a loading control²⁷. For protein expression experiments a Sma-1 fragment of DEC-205 from amino-acid position 8 to 633 (predicted M_r , 71.5K) was cloned in frame with GST (M_r , 26K; pGEX-2T, Pharmacia) and transformed into *E. coli* BL21. Western blotting with NLDC-145 monoclonal antibody was as previously described (W.J.S., M.A., M.C.N. and R.M.S., submitted). Symbols: Br, brain mRNA; DC, dendritic cell mRNA; Ht, heart mRNA; Kd, kidney mRNA; Lv, liver mRNA; LN, lymph node mRNA; Sk, skin mRNA; Th, thymus mRNA; Tg, tongue mRNA; GAP, glyceraldehyde-3-phosphate dehydrogenase; GST-DEC, *E. coli* lysates from cells transformed with the pGEX-2T-DEC-205 fusion construct; GST, lysates from *E. coli* cells transformed with pGEX-2T alone.

b

DEC-205 -27 MRTGRVTPGLAAGLILLLR-SFGLVFP
bvPLA2R -20 M-----PLLSSLLLLQLVPAGSA--
humfmar -18 MR-----LPILLVFASVIGAVL-
Lsignal sequence-

-----N1-----
1 -----SESSGND--PFTIVHNTGKCIQPLSDW-VVAQDCSGTNN-MLWKVWSQHRLFHLESQKCLGLDITKATDNLRM
1 ETAAMAVTPERLEWQDQKGFIIQSENLEKCIQASKST-LTLENCKPPNKVLMKWNHRLFNTGGSGCLGLNVSSPEQLPSI
1 -----LLDTQFLIYNEHDKRCVDAVSFAVQTAACQDAESQKFRVWSESQIMSVAFLCLGVPSKTDWVAITL
Lcysteine-rich domain

DEC-205 FSCDST-VMLWNKCEHHSLYTA-----AQYRLAKDGYA-VANTNTSDVWVK-GGSEENLCAQPYHEI
YECDSHVSILKHNCKKKTITGP-----LQYLQVQKQDNTLVASRYLHKVYMSYSGGGGICDYLHKDL
humfmar YACDSKSEFQKWECKNDTLGIGEDLFFNYGNRQEKINILYKSGSLWSRWKIY-GTIDNLCSRGYEAM

131 YTRDQNSYGRPCFPFLIGETWYHDCIHDE-DHSGPNCATTLSYE
147 YTIKGNHGTCPMPFPQYQNHHECTREGREDMLLWCATTSTRYE
139 YTLGLHANGATCAFFPKFNKNYADCTSAGRSQDGLWCGITTDYD
Lfibronectin type II domain

DEC-205 YDQKWCICLLP 186 ---ESG
bvPLA2R RDEKWCFCDDP 203 TSTEVG
humfmar TDKLFCYCPLK 196 FE---G
Lspacer 1

189 CEGNWEKNEQIGSCYQFNQELSWKRAYVSCQNGADILSIHSAELAYIT---GKEDIARLVNLGLNQLYSARGWESDPRPLKFLNWD
209 CDVWKKDLHSRICYQFNLLSSLSWSEHSSCQMGQALLSIADETEENFVRKHLGSEAVE---VWMLNQLDEADAGWQSDRTPNLNWLK
219 SESLWKKDLPLTSVSYQINSKALTWQARKSCQQNAELLSITEIHQTYLTGL---TSSLTSLGLNQLDLSLFSNMGQSDRSPFRLNWL
Lspacer 1 X ϕ 0 c 0 0 E X 0 ϕ 0 0 ϕ X g X X WZ

-----T52A-----
334 CHVGWLPNNGFCYLLAN-ESSSWDAHLKCKAFADLSHMSLADVE
358 CBPGNPNHNRCKLQK-EKKTWNEALQSCQSNVSLTDITSLAEVE
344 CPSQWNPYAGHCYKIRHDEKIQDALTTCRKGGLDTISITIELD
Lspacer 2 X ϕ 0 c 0 0 E

DEC-205 PGTPVAPVIGGSSCARMDT-ESGLWQSVSCSQPYVCKK
bvPLA2R PEINFPFVE-YHCOTFNAPMPKAWKSRDCESTLPYVCKK
humfmar PGFSABE---PGKSCVSLNPGKNAKWNLECVQKLGICCK
PBB E ϕ 0c0 X g WND c X c

316 ---PLNNTLELPDVTYTDTH
337 YLNPDPHGVVEKDKMYAETH
325 GNTTLNSFPVISE---SDVPTH
Lspacer 2

-----T52A-----
334 CHVGWLPNNGFCYLLAN-ESSSWDAHLKCKAFADLSHMSLADVE
358 CBPGNPNHNRCKLQK-EKKTWNEALQSCQSNVSLTDITSLAEVE
344 CPSQWNPYAGHCYKIRHDEKIQDALTTCRKGGLDTISITIELD
Lspacer 2 X ϕ 0 c 0 0 E

DEC-205 VVVKLHNGDVKKEIMTGLKNTNSPALFQWSDGTEVTLTYNNENEFVFPNKTNCVSYLQKLGQWQVSCCKKRYVCKK
bvPLA2R PLVT-LLGDENASETWIGLSSHKIPVSEWSSGVTFTNWHLEPHIFPNRSQLCVSAEQSGHVKVKNCEETFLYCKK
humfmar PIISQL-GYEPNDELWIGLNDIKIOMYFEWSDGTPVTFTKWLGEPSHENNRQEDCVVMKQDGYWADRGCEWPLGYICKM
X 0 ϕ 0 0 ϕ X g X X W ZPBB E ϕ 0c0 X g WND c X c

-----T16B-----
461 KGEITKDA---ESDKL 474 CPPDEGWKRRHE
484 TGLVLSDT---ESG-- 495 C---QKQWERRHK
471 KSRSGQPEIVEVEKG 486 CR---KQWKKRHF
Lspacer 3

DEC-205 TCYKIYEKAPFGTN-----CN-----L-TITSRFEQFLNMMKNYDKSLKRYFTGLRDPDSRGEYSNAVAQVQKQAVTSNNWFLEPASPGGCVAMSTGKTLQKWEVKNCRSFRALS
bvPLA2R FCYKIDTVLRSFHASSGYCPA-----LITITSRFEQAFITSLISSVVKTKDTYFNIALQDQNTGEYTWKTAGQQLEPVKYTHWTRQPRYSQGGCVMRGRSHPRGRWEVDRCDRHFAMK
humfmar YCYMIGHTLSTFAEANTQ---CNENAYLITIEDRYEQAFITSFVGL---RPEYFTNTGLSDITKGTQFWTIE---EEVRFTHWSDMRGKLPCCYAMRTGIAGGLMDVLTKCDE-KAKF
X ϕ 0 c 0 0 E X 0 ϕ 0 0 ϕ X g X X W ZPBB E ϕ 0c0 X g WND c X c

-----L68-----
461 KGEITKDA---ESDKL 474 CPPDEGWKRRHE
484 TGLVLSDT---ESG-- 495 C---QKQWERRHK
471 KSRSGQPEIVEVEKG 486 CR---KQWKKRHF
Lspacer 3

DEC-205 ICKK 599 -VSEPEEAAFPKDDP
bvPLA2R LCKQ 626 PVENRETKQEEGWFPHP
humfmar VCKH 610 WAEQVTHPPKPTTPEPK
c Lspacer 4

616 CPEGWHTFPSSLSYKVFHIERIVRKRNEEAERFCQALGAHLPSFRREEIKDPVH-LLKQDQS---GQRWNLNIGLNRKSRFDLQGS
644 CYLDNESEPLGLASCFKVFHSEKVLKRTWRQAEFCFEGQAHLASPAHIEE-ENFVNEILLHSPKNTREERQFNIQFNKRNPLNAGS
628 CPEDWGASSRTSLCFKLYAKKH-EKKTWNEALQSCQSNVSLTDITSLAEVE---QQTWILITA---SGSYHKLFWLGLTYGSP---SEG
Lspacer 4 X ϕ 0 c 0 0 E X 0 ϕ 0 0 ϕ X g X X W ZPBB E ϕ 0c0 X g WND c X c

-----L68-----
461 KGEITKDA---ESDKL 474 CPPDEGWKRRHE
484 TGLVLSDT---ESG-- 495 C---QKQWERRHK
471 KSRSGQPEIVEVEKG 486 CR---KQWKKRHF
Lspacer 3

DEC-205 QWSDRTFVSA---VMMEPEFQDDFDIRDCAAIKVLDPVRRVHLYEDKDIAYWKPACDAKLEWVCQI
bvPLA2R WEWSDQTPV-VS---SFLDNYFGE-DARMCAYKAN-----KTLT-----PSY-----CGSKREWICKI
humfmar FTWSDGSPSYENWAYGEPNNYQNV---EYCGELKGDPT-----MSWNDINCEHLNWCQI
 ϕ X g X X W ZPBB E ϕ 0c0 X g WND c X c

-----L73A-----
606 PKGSTQPMQDNYNPE---RTG 784 IHGPFVIEGSEIWEVAD
779 PRDVRFPKPPNYQ----- 792 YDAPWLFYQDAEYLFHS
762 QKQTEKPEPTAPQDNPV 782 TEDGWVYIKDYQYFES
Lspacer 5 X ϕ 0 c 0 0 E X 0 ϕ 0 0 ϕ X g X X W ZPBB E ϕ 0c0 X g WND c X c

DEC-205 PHLYEEAVLYCASNHSLATITSTGLKAIKKNLANISGEBQKRWVKTSENPIDRYFLGSRRLWHHPMTFG
bvPLA2R AS-EWSSSEFVCGWLRSDILTIRSAHEQEFIRSKIRALSRYGVNWNIGLREERASDEF---RWRDGSFVIYQWWDKGRSRMGLNESQRCGFISSITGLWAS---EBCSISMPSECKR
humfmar KE-TMDNARAFCKRNFQDLYSISQSESEKFLW-KYVNRNDAQSAFYIGLLIS---LDKKFAMMDGSKVDYVSWATGEFNT---ANEDENCVTMYSNSGFWND---INCGYPNAPFCQR
0 c 0 0 E X 0 ϕ 0 0 ϕ X g X X W ZPBB E ϕ 0c0 X g WND c X c

-----L73B-----
606 PKGSTQPMQDNYNPE---RTG 784 IHGPFVIEGSEIWEVAD
779 PRDVRFPKPPNYQ----- 792 YDAPWLFYQDAEYLFHS
762 QKQTEKPEPTAPQDNPV 782 TEDGWVYIKDYQYFES
Lspacer 5 X ϕ 0 c 0 0 E X 0 ϕ 0 0 ϕ X g X X W ZPBB E ϕ 0c0 X g WND c X c

DEC-205 907 YN---VSSLEKYSDFPAKQV
bvPLA2R 921 KK---VWVIEKKDKIPKQHT
humfmar 907 HSSINATTVMPTMPVSPSG
Lspacer 6

925 CTEKNIPFQNKCFLL-KVNSGQPVV---FSQASGICHSEYGGTFLSVLSRGEQDFIISLLPEMEASLNIGLWRTAYERINRMTNRELTYSNF
939 CPKQWLYFDYKCLLIRIPEGSPDMKNTSAQDFCVREGGTVAIERNEVEQAFITMNLGHTTNVHIGLQDDYK---MLNGRPVSYSNW
927 CKEGWNFYSNKKCFKI-POFMEERWQEAARKACIGFGQNLVSIQNEKQAFITMNMKDSFTSAWTLNVDNSEHTFMTDGRGVHYTNW
Lspacer 6 X ϕ 0 c 0 0 E X 0 ϕ 0 0 ϕ X g X X W ZPBB E ϕ 0c0 X g WND c X c

-----T60-----
606 PKGSTQPMQDNYNPE---RTG 784 IHGPFVIEGSEIWEVAD
779 PRDVRFPKPPNYQ----- 792 YDAPWLFYQDAEYLFHS
762 QKQTEKPEPTAPQDNPV 782 TEDGWVYIKDYQYFES
Lspacer 5 X ϕ 0 c 0 0 E X 0 ϕ 0 0 ϕ X g X X W ZPBB E ϕ 0c0 X g WND c X c

DEC-205 HPLLVGRRSLIPTNFFDDESH---FHCALILMLKKSPLTGTWNTFSCSERHSLSLCQR
bvPLA2R SP---PDTKNIPNNHTTEVQKRPLCGLLSNPNFHTQKWPEDCREGYGF-VCEK
humfmar GK-----GYPGRRSSLSYEDACVVIIGASNE-AGKMMDDTCDKRGY-ICQT
ZPBB E ϕ 0c0 X g WND c X c

-----T13-----
1066 YSETE 1071 DGQPHWNT---SKTVKYLNNLKIISKPLTHGALKCKMCKE
1079 MQDAS 1084 -GHSINTSDMYPINPTLEYGNRTYKIINAMTWYALTKCLMHG
1064 RSDPS 1069 LTNPPATIQ---TDGPFVYKGSYSILMRQKPFQWHEATYCKLHN
Lspacer 7 X ϕ 0 c 0 0 E X 0 ϕ 0 0 ϕ X g X X W ZPBB E ϕ 0c0 X g WND c X c

DEC-205 MRLVSTIDFYQQAFLAVQATLRNSSF---WIGLSSQDDELNFGWSDGKRLQFSNAGSNEQ-LDDCVILTDGFWKATDCDDNPGGAICYI
bvPLA2R AELASITDQYHQSLTV---ILNRYGVAHNIGLFTDNLGSLFSDSDGTSKSPFTFWKDGCVFADTSGRWSSTACSLYQGAICQV
humfmar SLIASILDFYSNAFAWLQ---METSNERVWIALNSMLTDNQYTWTDKWRVRYTNWADEPKLSACVYLLDGYWKTACHESFY-PLCKR
0 0 E X 0 ϕ 0 0 ϕ X g X X W ZPBB E ϕ 0c0 X g WND c X c

-----L17C-----
1197 PGNETEEVERALDTAK
1215 PTESTRLSGRLE----
1197 SDEIPATEPPQL-PGR
Lspacer 8

DEC-205 1213 CPSPVQSTPWIPFQNSCYNFMITNNRKTVTPEEVQSTCEKHLPKAHSLSIRNEBENTFVVEQLLYPNYIAS---WVMLGITYENNSLMWFDKTLASVTHWRTGRTFVKNKPLAG
bvPLA2R 1227 CSET---SIPWIKFKNCSYF---STVLETSFPAAEHFCCK---KGSMLITKDEAENSFLLELLAPRSSVQMIWLNQADGDNETIKWFDGTPDQSNWQIRKPEVYH---FKPH
humfmar 1212 CPES-DHTAWIPFHGHCY-YIESSY---TRNWQASLECLR---MGSSLSVIESAAESSFLSYRVEPLKSKTNF-WIGL-FRNVEGTWLMINNSPVSFVNMWTDGFSGER---N
Lspacer 8 X ϕ 0 c 0 0 E X 0 ϕ 0 0 ϕ X g X X W ZPBB E ϕ 0c0 X g WND c X c

-----L17C-----
1197 PGNETEEVERALDTAK
1215 PTESTRLSGRLE----
1197 SDEIPATEPPQL-PGR
Lspacer 8

DEC-205 LST-----DGFWDI-QSFNVIBETLHFYQHSISACKI
bvPLA2R LCVALRIPEGVQLSSCQDKKG-----FICKM
humfmar DCVALHASSGFWSNIHCSSYKG-----YICKR
0c0 X g WND c X c

1357 EM-VDYEDKNGH
1360 EADIHTVKKHFGK
1340 PKIIDAFTHELLTTKADTKMDPSK
Lspacer 9

-----L61A-----
1368 TLKQFIPYKDGVSIVQKKTWYALNACSGGELASVHNPN
Lspacer 9 X ϕ 0 c 0 0 E

DEC-205 GKLFLEDIVNRDGFPLWGLSSHDGSESSFMSDGRAPDYVPWQSLQSPGDCVLYPKGIWRREKCLSVKDGAIYK
bvPLA2R
humfmar
X 0 ϕ 0 0 ϕ X g X X W ZPBB E ϕ 0c0 X g WND c X c

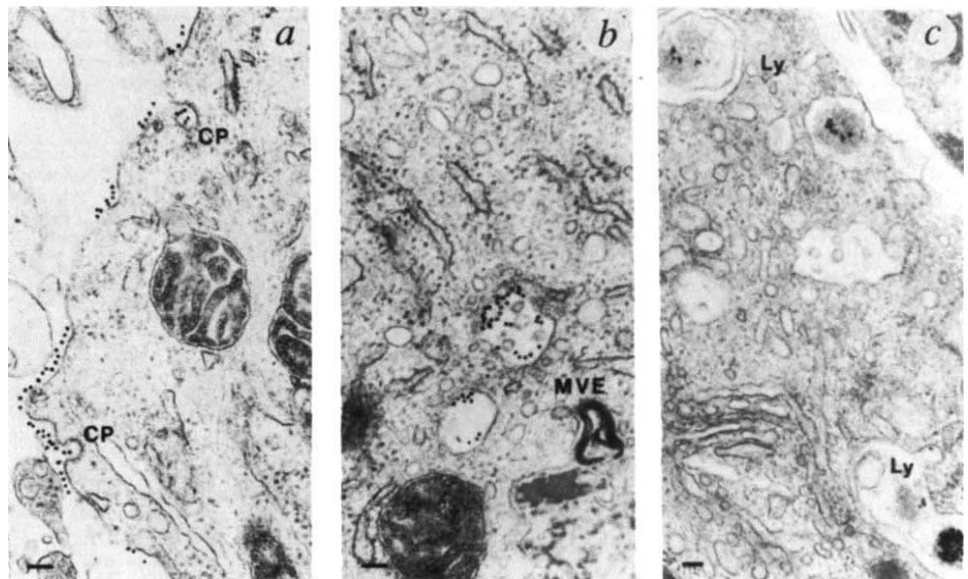
-----L60-----
1488 PTKDKKLIFHVKSSK 1503 CPVAKRDGPQWVQY
Lspacer 10

DEC-205 GGHCYASDQVLSHFSKQVQCELDHSATVVTIADENENKPVSRILKRENYITMRVWGLSQHSLSQSWLDGLDVTVPKWKNTKDGDKCSILIASNETWRKVCBSRGYARAVCKI
bvPLA2R
humfmar
X ϕ 0 c 0 0 E X 0 ϕ 0 0 ϕ X g X X W ZPBB E ϕ 0c0 X g WND c X c

DEC-205 1636 PLSPD 1641 -YTGIAI-LFAVLCILGLISLA-IWFL-
bvPLA2R 1374 GPHSVIPLTVALTLLVILAISTLSPCHY
humfmar 1366 PSSNVAGVVIIVILITGAGLAAYFFY-
Lspacer 11 Ltransmembrane domain

-----T27A-----
1666 QRSHI---RMTGFSVRYEHGTNEDEVMPLSFH---D- 1696
1402 KHSHIIFGLAQFRNPFYPSANFSTVHLEENILISDLKENDQ- 1443
1394 KRRVHLPOEGAFENTLYFNSQSSPGTSDMKDLVGNIEQNEHSVI 1438
Lcytosolic domain

FIG. 3 Endocytosis of DEC-205. Ultrastructural analysis of DEC-205 on dendritic cells with polyclonal rabbit anti-DEC-205 F(ab')₂ fragments and 10-nm gold-labelled goat anti-rabbit IgG (Amersham). Dendritic cells collected from 7-day mouse bone-marrow cultures²⁸ were incubated with 10 µg ml⁻¹ of either polyclonal anti-DEC-205 IgG, F(ab')₂ fragments of polyclonal anti-DEC-205, or biotinylated monoclonal NLDC-145, on ice for 30 min. Excess primary antibody was removed by washing cells 3 times with RPMI-1640, 10% FCS, 0.02% NaN₃. The cells were then incubated for 30 min on ice with either a 1:5 dilution of 10-nm gold-labelled goat anti-rabbit IgG, or a 1:5 dilution of 10-nm gold-labelled streptavidin, respectively. Excess secondary reagent was removed by washing cells as above. Dendritic cells were then either fixed with 2.5% glutaraldehyde for a time-zero point, or incubated for 1, 5, 20 or 60 min at 37 °C before fixation and processing for electron microscopy. For each time point, 10 grids were examined and all cells that were labelled with gold were photographed. Gold particles were counted and scored based on standard morphological criteria. The distribution of gold particles was as follows: 0 min, 158 (95%) on the plasma membrane (PM), 8 (5%) in coated pits or



coated vesicles (CP/V); 1 min, 68 (62%) PM, 31 (38%) CP/V; 5 min, 129 (41%) PM, 83 (28%) CP/V, 99 (32%) multivesicular endosome (MVE); 20 min, 12 (5%) PM, 3 (1%) CP/V, 175 (79%) MVE, 32 (14%) lysosomes (Ly); 60 min, 17 (7%) PM, 9 (5%) CP/V, 46 (19%) MVE, 172 (70%) Ly. a, 1 min; b, 20 min; c, 60 min. Scale bars, 100 µm.

205 (Fig. 1) suggested that this receptor might be used by dendritic cells and thymic epithelial cells to endocytose^{6,9,16,17} and deliver a variety of extracellular glycoprotein antigens to an intracellular antigen-processing compartment. To test this idea, we examined the fate of gold-labelled monoclonal and polyclonal anti-DEC-205 antibodies bound to DEC-205 on dendritic cells by electron microscopy in a time-course experiment (Fig. 3). At time zero, 95% of the cell-associated gold particles were found on the plasma membrane. After warming the sample to 37 °C for only one minute, 38% of the particles were found in coated vesicles or coated pits. By 20 minutes after crosslinking, 79% of the gold particles were in a multivesicular compartment that is

abundant in dendritic cells¹⁸ and resembles the MHC class II-containing vesicles that are thought to be involved in antigen presentation¹⁹⁻²³ (Fig. 3). Similar results were obtained in two separate experiments with monoclonal NLDC-145 antibody and intact or F(ab')₂ fragments of polyclonal anti-DEC-205 antibodies. Thus DEC-205 is rapidly taken up by means of coated vesicles, and antibodies bound to it are delivered to a morphologically distinct multivesicular endosomal compartment implicated in antigen processing.

To determine whether DEC-205 could deliver antigen to an active antigen-processing compartment, dendritic cells were treated with rabbit anti-DEC-205 antibodies and assayed for presentation of rabbit IgG-peptide-MHC complexes to T-cell hybridomas²⁴. Negative controls included non-immune rabbit IgG and rabbit antibodies to IgG2a that are efficiently presented by B-cell lines²⁴. Dendritic cells presented rabbit anti-DEC-205 to the T-cell clones two orders of magnitude more efficiently than the non-immune rabbit IgG or rabbit anti-IgG2a (Fig. 4). Thus DEC-205 resembles membrane immunoglobulin on B cells in that the crosslinked receptor is efficiently taken up, and bound ligands are delivered to an intracellular compartment that is active in antigen processing^{25,26}.

We conclude that dendritic cells and thymic epithelial cells express a novel receptor, DEC-205, which contributes to antigen presentation. The multi-lectin domain structure of this receptor suggests that it can be used by dendritic cells and thymic epithelial cells to capture and endocytose diverse carbohydrate-bearing antigens and direct them to the requisite compartments for antigen processing and presentation. □

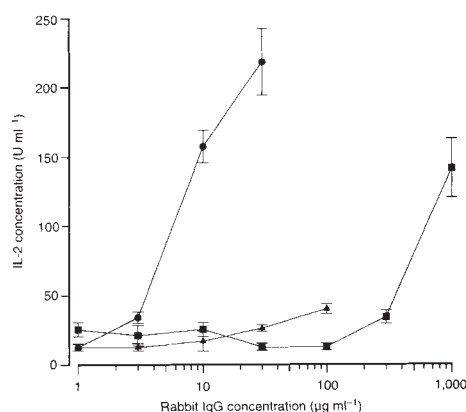


FIG. 4 Antigen presentation by dendritic cells incubated with rabbit anti-DEC-205 IgG or non-reactive rabbit IgG controls. BALB/c mouse dendritic cells (10⁵) obtained from day-7 bone-marrow cultures were cocultured with 10⁵ 2R.50 rabbit IgG-specific T-hybridoma cells for 24 h in triplicate²⁴. The supernatants were assayed for IL-2 concentration using the HT-2 indicator cell line. IL-2 production by the 2R.50 cells is plotted against the concentration of antibody in the cultures on a logarithmic scale. Error bars indicate standard deviation from the mean. Symbols: circles, cultures that received the indicated amounts of polyclonal rabbit anti-DEC-205 IgG; triangles, cultures that received the indicated amounts of IgG2a-specific polyclonal rabbit antibodies; squares, cultures that received the indicated amounts of nonspecific rabbit IgG.

Received 4 January; accepted 8 March 1995.

- Steinman, R. M. A. *Rev. Immun.* **9**, 271-296 (1991).
- Matzinger, P. & Gunder, S. *Nature* **338**, 74-76 (1989).
- Hugo, P., Kappler, J. W. & Marrack, P. C. *Immunol. Rev.* **135**, 133-155 (1993).
- Kraal, G., Bree, M., Janse, M. & Bruin, G. J. *exp. Med.* **163**, 981-997 (1986).
- Taylor, M. E., Conary, J. T., Lennartz, M. R., Stahl, P. D. & Drickamer, K. *J. biol. Chem.* **265**, 12156-12162 (1990).
- Ezekowitz, R. A. B., Sastry, K., Bailly, P. & Warner, A. J. *exp. Med.* **172**, 1785-1794 (1990).
- Harris, N., Super, M., Rits, M., Chang, G. & Ezekowitz, R. A. *Blood* **80**, 2363-2373 (1992).
- Lambeau, G., Ancian, P., Barhanin, J. & Lazdunski, M. *J. biol. Chem.* **269**, 1575-1578 (1994).
- Ishizaki, J. et al. *J. biol. Chem.* **269**, 5897-5904 (1994).
- Drickamer, K. & Taylor, M. E. A. *Rev. Cell Biol.* **9**, 237-264 (1993).

11. Taylor, M. E., Bezouska, K. & Drickamer, K. *J. biol. Chem.* **267**, 1719–1726 (1992).
12. Drickamer, K. *Nature* **360**, 183–186 (1992).
13. Bezouska, K. *et al. Nature* **372**, 150–157 (1994).
14. Daniels, B. F., Nakamura, M. C., Rosen, S. D., Yokoyama, W. M. & Seaman, W. E. *Immunity* **1**, 785–792 (1994).
15. Taylor, M. E. & Drickamer, K. *J. biol. Chem.* **268**, 399–404 (1993).
16. Chen, W.-J., Goldstein, J. L. & Brown, M. S. *J. biol. Chem.* **265**, 3116–3123 (1990).
17. Collawn, J. F. *et al. Cell* **63**, 1061–1072 (1990).
18. Kleijmeer, M. J., Oorschot, V. M. J. & Geuze, H. J. *J. invest. Derm.* **103**, 516–523 (1994).
19. Peters, P. J., Neefjes, J. J., Oorschot, V., Ploegh, H. L. & Geuze, H. J. *Nature* **349**, 669–676 (1991).
20. Amigorena, S., Drake, J. R., Webster, P. & Mellman, I. *Nature* **369**, 113–120 (1994).
21. West, M. A., Lucocq, J. M. & Watts, C. *Nature* **369**, 147–151 (1994).
22. Tulp, A., Verwoerd, D., Dobberstein, B., L., P.H. & Pieters, J. *Nature* **369**, 120–126 (1994).
23. Qiu, Y., Xu, X., Wandinger-Ness, A., Dalke, D. P. & Pierce, S. K. *J. Cell Biol.* **125**, 595–609 (1994).
24. Boom, W. H., Liano, D. & Abbas, A. K. *J. exp. Med.* **167**, 1350–1364 (1988).
25. Rock, K. L., Benacerraf, B. & Abbas, A. K. *J. exp. Med.* **160**, 1102–1125 (1984).
26. Lanzavecchia, A. *Nature* **314**, 538–539 (1985).
27. Piechaczyk, M. *et al. Nucleic Acids Res.* **12**, 6951–6963 (1984).
28. Inaba, K. *et al. J. exp. Med.* **176**, 1693–1702 (1992).

ACKNOWLEDGEMENTS. W.J. and W.J.S. contributed equally to this work. We thank Y. Choi for help with IL-2 assays, R. Mitchell for valuable advice and 2.R50 cells, Z. Misulovin and J. Meltzer for technical assistance, S. Gezelter and J. Adams for expert help with computer graphics, and H. Shio for assistance with electron microscopy. The Laboratory of Molecular Immunology is part of the Howard Hughes Medical Institute. This work was supported in part by the NIH (R.M.S.). W.J.S. is supported by a Young Scientist Training Program Scholarship from the Life and Health Insurance Medical Research Fund, and C.H. was supported by the Max Kade Foundation.

Natural and synthetic non-peptide antigens recognized by human $\gamma\delta$ T cells

Yoshimasa Tanaka^{*§}, Craig T. Morita[†],
Yoko Tanaka^{*§}, Edward Nieves[†],
Michael B. Brenner[†] & Barry R. Bloom^{*}

^{*} Howard Hughes Medical Institute and Departments of Microbiology and Immunology and [†] Biochemistry, Albert Einstein College of Medicine, Bronx, New York 10461, USA

[†] Lymphocyte Biology Section, Department of Rheumatology and Immunology, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts 02115, USA

T LYMPHOCYTES express either $\alpha\beta$ or $\gamma\delta$ T-cell receptor heterodimers^{1,2}. Most $\alpha\beta$ T cells recognize antigenic peptides bound to major histocompatibility complex molecules but the antigen recognition and biological function of $\gamma\delta$ T cells is unknown. A major human $\gamma\delta$ T-cell subset expressing V γ 2 and V δ 2 germline genes, but having diverse junctional sequences, is found in human mycobacterial lesions³ and responds *in vitro* to antigens of bacteria and parasites^{4–8}. In addition, certain haematopoietic tumour cells are specifically recognized and lysed by these T cells⁹. V γ 2V δ 2-bearing T cells were shown to recognize mycobacterial antigens that are protease resistant and phosphatase sensitive^{10–13}. Because of the difficulty in isolating natural antigens from mycobacterial culture filtrates or extracts, we synthesized a series of monoalkyl phosphates, and found that some, particularly monoethyl phosphate, could mimic the activity of mycobacterial antigens in stimulating these $\gamma\delta$ T cells¹⁰. Here we report the identification of natural antigens produced by mycobacteria recognized by human V γ 2V δ 2-bearing T cells as isopentenyl pyrophosphate and related prenyl pyrophosphate derivatives, compounds involved in the synthesis of complex polyisoprenoid compounds in microbial and mammalian cells. Substitution of phosphate for the pyrophosphate moiety, or elimination of the double bond, greatly reduced antigenic activity of these compounds. These results provide formal evidence that, in contrast to recognition of major histocompatibility complex-bound peptide antigens by $\alpha\beta$ T cells, human $\gamma\delta$ T cells can recognize naturally occurring small non-peptidic antigens.

To characterize the backbone structure of the natural mycobacterial antigens recognized by human V γ 2V δ 2-bearing T cells, mycobacterial preparations were carefully examined by adsorption and anion-exchange column chromatography. As previously described, activated charcoal failed to adsorb the vast majority of mycobacterial antigens recognized by $\gamma\delta$ T cells in the presence of high salt, but most chemical moieties including these antigens remained on the column under low salt concentrations (data not shown). This suggested that the microbial antigens might have carbon-carbon double bonds. On anion-exchange column chromatography, the natural antigens from *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* (a rapidly growing mycobacterium) eluted between the known elution peaks of inorganic phosphate and pyrophosphate (Fig. 1a, b), where most pyrophosphomonoesters would be expected to elute, with minor variations in retention. This suggested that the natural antigens contained a carbon-carbon double bond and a pyrophosphate moiety.

To determine the structure of the natural mycobacterial antigens, the active fractions in Fig. 1b were examined by parent ion scan of 79 AMU (atomic mass units) for phosphate (PO₃⁻) using electrospray ionization (ESI) tandem mass spectrometry (MS/MS) in the negative ion mode. The three major compounds in the active fraction (Fig. 1c), the 245 *m/z* (ionic mass) peak which is the most abundant, the 275 peak, and the 339 peak, were structurally deduced after fragmentation analysis by ESI-MS/MS. The 245 ion peak was identified as isopentenyl pyrophosphate, because the daughter ion spectra obtained from the 245 ion of *M. smegmatis* preparations were identical to those of the 245 molecular ion of synthetic isopentenyl pyrophosphate (Fig. 1d, e). The 275 peak was similarly identified as its hydroxymethyl derivative. The 339 peak is a hexose diphosphate, which was immunologically inactive. To confirm that isopentenyl pyrophosphate is a bioactive molecule, we synthesized and tested the ability of this compound to stimulate V γ 2V δ 2-bearing T cells. Although isopentenyl pyrophosphate, as well as several related prenyl pyrophosphate compounds, stimulated proliferation of $\gamma\delta$ T cells, isomyl pyrophosphate, a structural analogue of isopentenyl pyrophosphate lacking a carbon-carbon double bond, had no activity at the same range of antigen concentrations (Fig. 2a). To assess the biological activity of the 275 ion, isopentenyl pyrophosphate was selectively removed from the purified material through repeated lyophilization from dilute ammonium acetate and the hydrolysate was analysed for phosphate (PO₃⁻) by ESI-MS/MS as shown in Fig. 2b. In preparations from which the 245 ion species was depleted, the 275 ion species was found to be active (data not shown). This is further supported by the fact that the 275 ion represents the major pyrophosphorylated species in a highly purified non-peptidic antigen preparation from *Mycobacterium fortuitum* (data not shown). We have synthesized 6- and 7-carbon derivatives of dimethylallyl pyrophosphate and found them to be 30- to 50-fold more active than the parent compound, suggesting that the hydroxymethyl derivative may be more active than isopentenyl pyrophosphate itself. These results unambiguously establish that isopentenyl pyrophosphate and related prenyl pyrophosphate derivatives exist in mycobacteria as natural antigens for human V γ 2V δ 2-bearing T cells. To establish the involvement of the T-cell receptor (TCR) in the recognition of the natural non-peptide antigens by this subset of T cells, we examined the response of mutant Jurkat cells, a human T-cell line lacking TCR molecules, transfected with complementary DNAs encoding various $\gamma\delta$ TCRs, as described previously¹⁴ for their ability to respond to isopentenyl pyrophosphate. Although the transfectants only express $\gamma\delta$ TCRs at a level of one-tenth that of normal human $\gamma\delta$ T-cell clones, mutant Jurkat cells transfected with cDNAs encoding a V γ 2V δ 2 TCR recognized isopentenyl pyrophosphate in a dose-dependent manner, but not the biologically inactive compound, phenethyl pyrophosphate (Fig. 2c, right panel). In contrast, mutant Jurkat cells transfected with cDNAs encoding

§ Present address: Osaka Bioscience Institute, 6-2-4 Furuedai, Suita, Osaka 565, Japan.

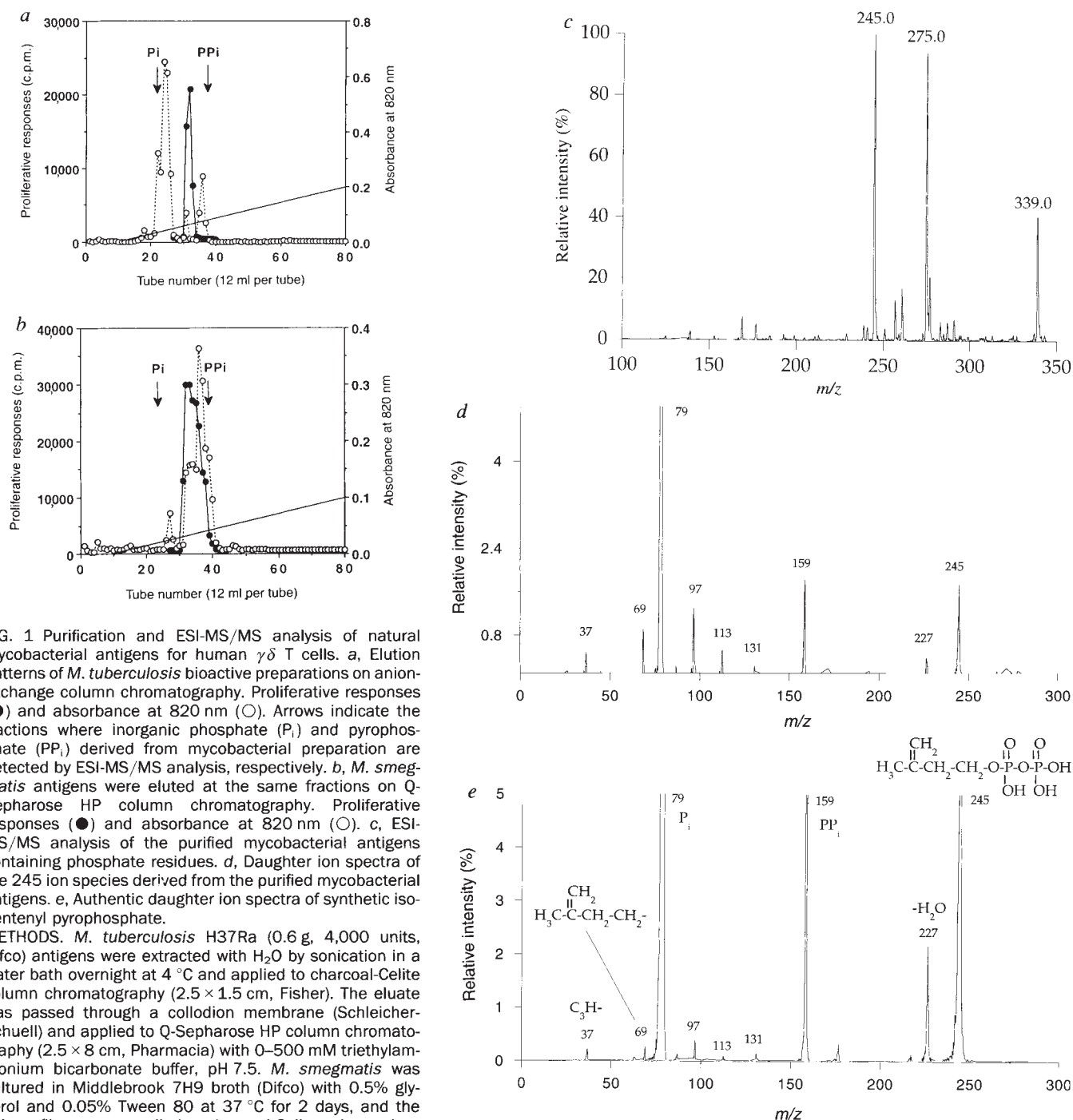


FIG. 1 Purification and ESI-MS/MS analysis of natural mycobacterial antigens for human $\gamma\delta$ T cells. *a*, Elution patterns of *M. tuberculosis* bioactive preparations on anion-exchange column chromatography. Proliferative responses (●) and absorbance at 820 nm (○). Arrows indicate the fractions where inorganic phosphate (P_i) and pyrophosphate (PP_i) derived from mycobacterial preparation are detected by ESI-MS/MS analysis, respectively. *b*, *M. smegmatis* antigens were eluted at the same fractions on Q-Sepharose HP column chromatography. Proliferative responses (●) and absorbance at 820 nm (○). *c*, ESI-MS/MS analysis of the purified mycobacterial antigens containing phosphate residues. *d*, Daughter ion spectra of the 245 ion species derived from the purified mycobacterial antigens. *e*, Authentic daughter ion spectra of synthetic isopentenyl pyrophosphate.

METHODS. *M. tuberculosis* H37Ra (0.6 g, 4,000 units, Difco) antigens were extracted with H_2O by sonication in a water bath overnight at 4 °C and applied to charcoal-Celite column chromatography (2.5 × 1.5 cm, Fisher). The eluate was passed through a collodion membrane (Schleicher-Schuell) and applied to Q-Sepharose HP column chromatography (2.5 × 8 cm, Pharmacia) with 0–500 mM triethylammonium bicarbonate buffer, pH 7.5. *M. smegmatis* was cultured in Middlebrook 7H9 broth (Difco) with 0.5% glycerol and 0.05% Tween 80 at 37 °C for 2 days, and the culture filtrate was applied to charcoal-Celite column chromatography (2.5 × 1.5 cm). The mycobacteria were again cultured in the clear eluate supplemented with 7H9, glycerol and Tween 80 under the same condition. This procedure was repeated 9 times and 7 million units were obtained. One unit is defined as the antigen concentration required for half-maximal proliferative responses of $\gamma\delta$ T-cell clone 12G12, derived from the lymphocytes of a tuberculous leprosy patient¹⁵. To the solution, 5 vols of methanol were added, and the clear supernatant was evaporated to dryness under reduced pressure at room temperature. The residue was rehydrated in 1 l H_2O and passed through a collodion membrane. The filtrate was aliquoted into 40 portions and applied to Q-Sepharose HP column chromatography (2.5 × 8 cm) with between 0 and 500 mM triethylammonium bicarbonate buffer, pH 7.5. Active fractions from 40 column chromatography runs were combined and applied to cellulose (Sigma) flash column chromatography with isopropanol:concentrated $NH_4OH:H_2O$ (7:2:1, v/v/v) to remove nucleotide contaminants, then the active fractions were rechromatographed by Q-Sepharose HP (2.5 × 8 cm) as shown in *b*. For phosphate assay²⁵, 10- μ l samples were used in *a* and *b*, and 10 or 0.01 μ l samples were used for proliferation assay, respectively. Proliferation assays were

done in triplicate using 5×10^4 cloned $\gamma\delta$ T cells per round-bottom well of a 96-well plate. After 36 h, the cells were pulsed with [methyl- 3H]thymidine (1 μ Ci per well), collected at 48 h, and counted by liquid scintillation. Mass spectral analysis was done on a API-III triple-quadrupole mass spectrometer (PE-SCIEX, Ontario, Canada) using the SCIEX-IonSpray interface with nitrogen as the nebulizer gas. Flow injection analysis using 100% methanol plus 0.1% trifluoroacetic acid at 10 μ l min^{-1} was used to introduce the samples into the mass spectrometer. Collisional activated dissociation (CAD) was done using argon as the collision gas at a collision gas thickness of 1.7×10^{10} molecules cm^{-2} and a collision energy of 30 eV. Under these CAD conditions, parent ion scans of 79 AMU and daughter ion spectra of the selected parent ion, m/z 245, were obtained. Samples were diluted 1:50 using 0.2% NH_4OH and a 10 μ l aliquot was injected. Of the three peaks in *c*, the 339 ion was identified as a hexose diphosphate (immunologically inactive), and the 245 and 275 ions were identified as isopentenyl pyrophosphate and its hydroxymethyl derivative, possibly derived from 5-diphosphomevalonic acid, respectively, based on the daughter ion analysis (data not shown) and synthetic data.

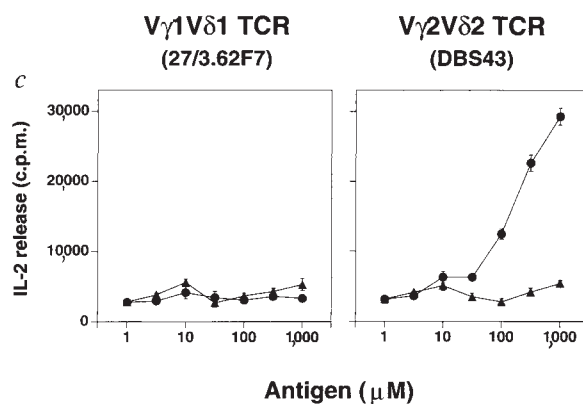
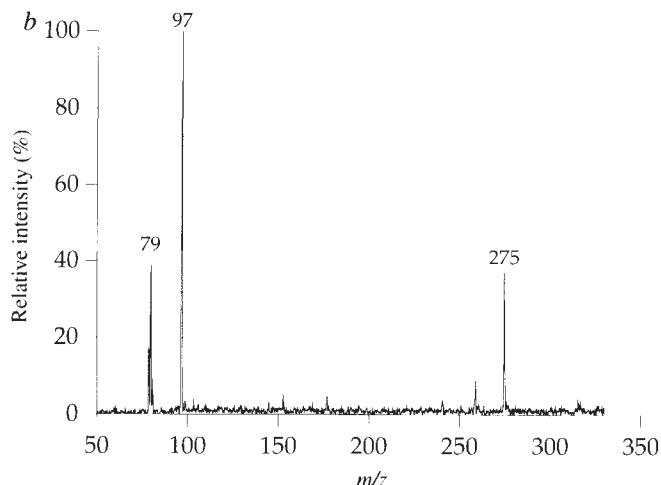
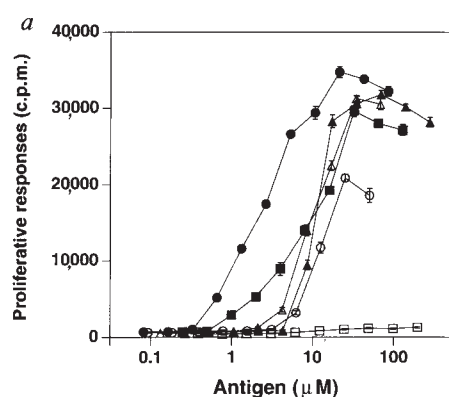


FIG. 2 TCR-dependent recognition of prenyl pyrophosphate derivatives by $\gamma\delta$ T-cell clone 12G12 and by the $V\gamma 2V\delta 2$ TCR transfectant, DBS43. **a**, Proliferation of 12G12 in response to isoprene units coupled with pyrophosphate. Pyrophosphate derivatives of isopentenol (●), dimethylallyl alcohol (Δ), geraniol (▲), farnesol (■), geranylgeraniol (○) and isoamyl alcohol (□). **b**, Parent ion scan for phosphate (PO_3^-) of a purified mycobacterial preparation depleted of isopentenyl pyrophosphate but with biological activity. **c**, TCR-dependent recognition of isopentenyl pyrophosphate by the $V\gamma 2V\delta 2$ TCR transfectant, DBS43. Interleukin-2 (IL-2) release by the $V\gamma 2V\delta 2$ TCR transfectant, DBS43 (right panel), or by the control $V\gamma 1V\delta 1$ TCR transfectant, 27/3.62F7 (left panel), in response to isopentenyl pyrophosphate (●) or to the biologically inactive compound, phenethyl pyrophosphate (▲). **METHODS**. Prenyl, isoamyl and phenethyl pyrophosphates were from Sigma or prepared by reacting ditriethylammonium phosphate with the indicated alcohol in the presence of trichloroacetonitrile²⁴, and purified by Q-Sepharose HP column chromatography (2.5 × 8 cm) with 0–500 mM triethylammonium bicarbonate buffer, pH 7.5. The products were eluted at 180 mM. Proliferation assays were as described in Fig. 1. For the hydrolysis of isopentenyl pyrophosphate, the purified preparation was dissolved in 200 mM ammonium acetate and lyophilized three times in **b**. Two peaks, 79 (PO_3^-) and 97 ($H_2PO_4^-$) represent inorganic phosphate released by hydrolysis. For determining transfectant recognition, 1×10^5 cells of the $V\gamma 2V\delta 2$ TCR transfectant, DBS43, or the control $V\gamma 1V\delta 1$ TCR transfectant, 27/3.62F7, were stimulated in the presence of 2×10^5 DG.EBV cells in 0.2 ml media with 10 ng ml^{-1} TPA. To triplicate wells of a 96-well plate containing the cells was added either isopentenyl pyrophosphate, phenethyl pyrophosphate (negative

a $V\gamma 1V\delta 1$ TCR recognized neither isopentenyl pyrophosphate nor phenethyl pyrophosphate (Fig. 2c, left panel). Thus, the recognition of isopentenyl pyrophosphate by $\gamma\delta$ T cells is both TCR dependent and antigen specific.

It has been reported that a γ -derivative of thymidine triphosphate (TTP) had stimulating activity for $V\gamma 2V\delta 2$ -bearing

TABLE 1 Synthetic alkenyl and prenyl derivatives of phosphate, pyrophosphate and UTP stimulate human $V\gamma 2V\delta 2$ -bearing T cells

Anchor chain	Structure	Derivative (μM)*		
		Phosphate	Pyrophosphate	UTP
Alkenyl-				
Allyl-	$\text{CH}_2=\text{CHCH}_2-$	800	5	4
Crotyl-	$\text{CH}_3\text{CH}=\text{CHCH}_2-$	350	8	3
Prenyl-				
Dimethylallyl-	$\begin{array}{c} \text{CH}_3 \\ \\ \text{CH}_3\text{C}=\text{CHCH}_2- \end{array}$	30	10	0.3
Isopentenyl-	$\begin{array}{c} \text{CH}_3 \\ \\ \text{CH}_2=\text{CCH}_2\text{CH}_2- \end{array}$	700	3	4

* Values given are the antigen concentration (μM) required for half-maximal proliferative responses of $\gamma\delta$ T-cell clone 12G12. **METHODS**. Alkenyl and prenyl derivatives of phosphate and pyrophosphate were prepared by allowing ditriethylammonium phosphate to react with the indicated alcohol in the presence of trichloroacetonitrile²⁴. UTP γ -derivatives were prepared as described in Fig. 3. The products were chromatographed on a Q-Sepharose high-performance column (2.5 × 8 cm) with 0–500 mM triethylammonium bicarbonate buffer, pH 7.5 and eluted at 50, 180 and 280 mM, respectively. Proliferation assays were as described in Fig. 1.

control), or anti-TCR $\delta 1$ monoclonal antibody (1/2,000 dilution of ascites, positive control). After 24 h, the supernatants were collected, diluted 1/8, and used in an IL-2 assay with CTLL-20 indicator cells as described¹⁴. Control IL-2 release in the presence of anti-TCR $\delta 1$ antibody for the DBS43 transfectant was $24,271 \pm 1,452$ whereas IL-2 release for the 27/3.62F7 transfectant was $20,873 \pm 2,973$.

T cells, although the chemical nature of the antigen was not established¹³. On the basis of this hypothesis and our previous synthetic results¹⁰, a γ -monoethyl derivative of thymidine triphosphate (Et-TTP) was synthesized and tested for stimulation of a $\gamma\delta$ T-cell clone. As shown in Fig. 3a, the $\gamma\delta$ T cells proliferated in a dose-dependent manner to Et-TTP, reaching half-maximal proliferation at 3 μM , but failed to proliferate when exposed to monoethyl derivatives of thymidine di- or monophosphate. In addition to Et-TTP, all γ -monoethyl derivatives of nucleoside and deoxynucleoside triphosphate (NTP and dNTP) induced proliferation in $\gamma\delta$ T cells to different degrees (Fig. 3a, b), with uridine triphosphate (UTP) and TTP γ -derivatives exhibiting the highest specific activity. To characterize further the functional specificity of antigen recognition by $V\gamma 2V\delta 2$ -bearing T cells, we synthesized and tested a variety of small alkenyl and prenyl derivatives of phosphate, pyrophosphate and UTP. As shown in Table 1, derivatives of pyrophosphate and UTP had specific activities 50–250 times greater than those of the corresponding phosphomonoesters. For example, the half-maximal stimulating activity of the isopentenyl derivatives of phosphate, pyrophosphate and UTP were 700, 3 and 4 μM , respectively, suggesting that a pyrophosphate moiety is required for high specific recognition of the antigens by $\gamma\delta$ T cells. The responsiveness to the non-nucleotide derivatives indicates that $\gamma\delta$ T cells predominantly recognize the pyrophosphomonoester moieties of NTP or dNTP γ -derivatives, but not the nucleoside or deoxynucleoside residues. We speculate that the pyrophosphomonoesters may play a more biologically important role than

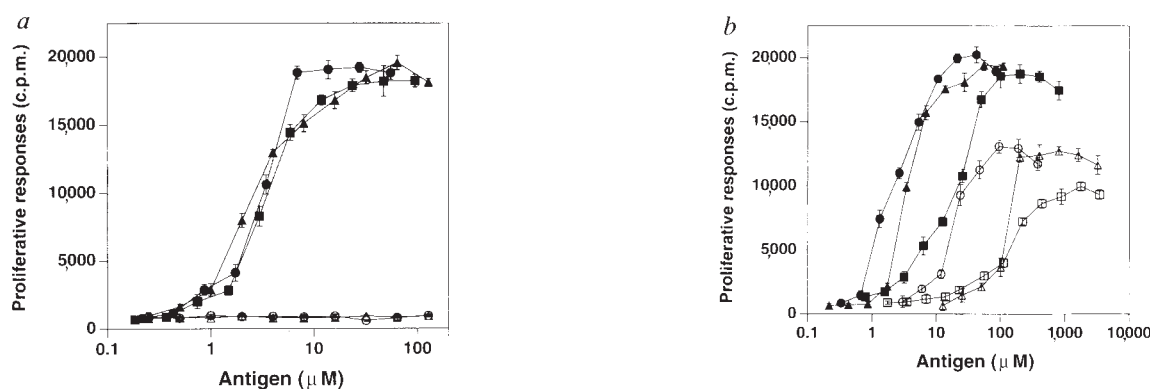


FIG. 3 Recognition of NTP or dNTP γ -derivatives by human $\gamma\delta$ T-cell clone 12G12 and a schematic diagram of isoprenoid pathways in bacterial and mammalian cells. **a**, Stimulation by dNTP γ -monoethyl derivatives. γ -Monoethyl derivatives of: 2'-deoxythymidine triphosphate (2'-dTTP, $\bullet$); 2',3'-dideoxythymidine triphosphate (2',3'-ddTTP, $\blacktriangle$); 2'-deoxyuridine triphosphate (2'-dUTP, $\blacksquare$); 2'-deoxythymidine diphosphate (2'-dUDP, $\circ$); 2'-deoxythymidine monophosphate (2'-dUMP, $\triangle$). **b**, Stimulation by NTP γ -monoethyl derivatives. γ -Monoethyl derivatives of: uridine triphosphate (UTP, $\bullet$); xanthosine triphosphate (XTP, $\blacktriangle$); inosine triphosphate (ITP, $\blacksquare$); guanosine triphosphate (GTP, $\circ$); cytosine triphosphate (CTP, $\triangle$); adenosine triphosphate (ATP, $\square$). **c**, Schematic diagram of isoprenoid biosynthetic pathways, and the relationship between the intermediates and oncogene modifications.

METHODS. NTP and dNTP γ -derivatives were prepared by allowing the triethylammonium salt of the indicated NTP or dNTP to react with the corresponding alcohol^{26,27}, and purified by anion-exchange chromatography on a Q-Sepharose HP column (2.5 $\times$ 8 cm) with 0–500 mM triethylammonium bicarbonate buffer, pH 7.5. The products were eluted at 280 mM. Proliferation assays were as described in Fig. 1. The mean and s.e.m. are given. In the schematic diagram, PP represents a pyrophosphate residue.

the corresponding NTP or dNTP γ -derivatives in $\gamma\delta$ T-cell activation, because the latter compounds remain predominantly in the cytoplasmic compartment of mycobacteria, whereas the pyrophosphomonoesters are secreted. In the case of *M. smegmatis* cultures, 98% of the total antigenic activity for $\gamma\delta$ T cells was found in the culture supernates. The nucleotide derivatives may have greater specific activity because of their interaction with cell membranes, or because they protect the essential pyrophosphate moiety from enzymatic degradation. Although the relative biological importance of nucleotide-derivatives and pyrophosphomonoesters remains to be determined, our results suggest that there may be a number of phosphorylated non-peptidic compounds produced by mycobacteria that stimulate human V δ 2V γ 2-bearing T cells.

Isopentenyl and related prenyl pyrophosphate derivatives are derived from acetyl coenzyme A and known to be the activated precursors of important lipophilic compounds such as vitamins, lipids and sterols as shown in Fig. 3c (refs 15–18). The recognition of prenyl pyrophosphate derivatives by the V γ 2V δ 2 subset of human $\gamma\delta$ T cells may explain their broad reactivity to a number of pathogens. Many bacteria can convert prenyl pyrophosphates to polyprenols, which serve as carriers for transport of mycolic acids across the mycobacterial cell wall¹⁹, and may be important as carriers of carbohydrate components to bacterial cell walls²⁰. In mammalian cells, farnesyl addition appears to be a critical modification for the membrane association of the *ras* protein and is required for transforming activity^{15, 18}. This common set of prenyl pyrophosphate intermediates, present in both bacterial and mammalian cells, may represent a common link in recognition by V γ 2V δ 2-bearing T cells of both microbial pathogens and haematopoietic tumour cells. The increased accumulation of $\gamma\delta$ T cells in lesions of human mycobacterial infection³, and their ability to respond to virus²¹ and bacterially infected cells *in vitro*, suggest that their response to such antigens could contribute both to pathogenesis and protective immune responses. Our preliminary data indicate that mice with a disruption

in the gene for TCR δ -chain die from *M. tuberculosis* infection whereas intact controls do not (K. Triebold, R. Mazzaccaro, J. Flynn and B.R.B., manuscript in preparation). The recognition of prenyl pyrophosphate derivatives by human $\gamma\delta$ T cells and mycolic acid by CD4⁺CD8[−] $\alpha\beta$ T cells²³, argues that alternative systems of specific T-cell recognition, complementing those for major histocompatibility complex-presented conventional peptidic antigens and superantigens, exist for non-peptidic antigens. □

Received 1 November 1994; accepted 15 March 1995.

- Marrack, P. & Kappler, J. *Adv. Immun.* **38**, 1–30 (1986).
- Porcelli, S., Brenner, B. B. & Band, H. *Immun. Rev.* **120**, 137–183 (1991).
- Modlin, R. L. et al. *Nature* **339**, 544–548 (1989).
- Kabelitz, D. et al. *J. exp. Med.* **171**, 667–679 (1990).
- Panchamoorthy, G. et al. *J. Immun.* **147**, 3360–3369 (1991).
- De Libero, G. et al. *J. exp. med.* **173**, 1311–1322 (1991).
- Hara, T. et al. *J. clin. Invest.* **90**, 204–210 (1992).
- Goerlich, R. et al. *Eur. J. Immun.* **21**, 2613–2616 (1991).
- Fisch, P. et al. *Science* **250**, 1269–1273 (1990).
- Tanaka, Y. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 8175–8179 (1994).
- Pfeffer, K. et al. *Eur. J. Immun.* **20**, 1175–1179 (1990).
- Schoel, B., Sprenger, S. & Kaufmann, S. H. E. *Eur. J. Immun.* **24**, 1886–1892 (1994).
- Constant, P. et al. *Science* **264**, 267–270 (1994).
- Bukowski, J. F. et al. *J. Immun.* **154**, 998–1006 (1995).
- Khosravi, F. R., Cox, A. D., Kato, K. & Der, C. J. *Cell Growth Differ.* **3**, 461–469 (1992).
- Goldstein, J. L. & Brown, M. S. *Nature* **343**, 425–430 (1990).
- Hancock, J. F., Magee, A. I., Childs, L. E. & Marshall, C. J. *Cell* **57**, 1167–1177 (1989).
- Schafer, W. R. et al. *Science* **245**, 379–385 (1989).
- Besra, G. S. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 12735–12739 (1994).
- Myant, N. B. *The Biology of Cholesterol and Related Steroids* 123–225 (Heinemann, London, 1981).
- Bukowski, J. F., Morita, C. T. & Brenner, M. B. *J. Immun.* **153**, 5133–5140 (1994).
- Havlic, D. V., Ellner, J. J., Chervenak, K. A. & Boom, W. H. *J. clin. Invest.* **87**, 729–733 (1991).
- Beckman, E. M. et al. *Nature* **372**, 691–694 (1994).
- Popjak, G. et al. *J. biol. Chem.* **237**, 56–61 (1962).
- Ames, B. N. *Meth. Enzym.* **8**, 115–118 (1966).
- Eckstein, F., Bruns, W. & Parmeggiani, A. *Biochemistry* **23**, 5225–5232 (1975).
- Knorre, D. G., Kurbatov, V. A. & Samukov, V. V. *FEBS Lett.* **70**, 105–108 (1976).

ACKNOWLEDGEMENTS. We thank E. Hehre, J. Blanchard, S. Sano, J. Bukowski, R. Modlin, G. De Libero and X. Tang for helpful discussion, and the Laboratory for Macromolecular Analysis at AECOM for the ES mass spectrometry facility supported by the NIH. This research is supported by research grants from the Howard Hughes Medical Institute (B.R.B., Y.T.), the NIH (B.R.B., M.B.B.), the Leukemia Society Scholar Award (M.B.B.) and a fellowship from the Arthritis Foundation (C.T.M.).

Cell-cycle inhibition by independent CDK and PCNA binding domains in p21^{Cip1}

Yan Luo, Jerard Hurwitz & Joan Massagué*

Cell Biology and Genetics Program, Molecular Biology Program and Howard Hughes Medical Institute, Memorial Sloan Kettering Cancer Center, New York, New York 10021, USA

* To whom correspondence should be addressed

MAMMALIAN cell-cycle control by antimitogenic signals involves p21^{Cip1/WAF1} (refs 1–4), p27^{Kip1} (refs 5, 6) and p57^{Kip2} (refs 7, 8), a family of proteins that bind to and inhibit cyclin-dependent kinases (CDKs) required for initiation of S phase. The protein p21 also binds to the DNA polymerase δ processivity factor, proliferating-cell nuclear antigen (PCNA), and inhibits *in vitro* PCNA-dependent DNA replication^{9,10}. The CDK and PCNA inhibitory activities of p21 are shown here to be functionally independent and to reside in separate protein domains. The PCNA binding and inhibitory activities, which are not observed with p27 or p57, reside in the C-terminal domain of p21, whereas the CDK inhibitory activity resides in the conserved N-terminal domains of these proteins. When separately overexpressed in mammalian cells, the CDK and PCNA inhibitory domains prevent DNA replication, demonstrating a dual function of p21 as a cell-cycle inhibitor *in vivo*.

The presence of CDK and PCNA inhibitory activities in p21 raises questions about the interdependence of these activities, their conservation in p27 and p57, and in particular their ability individually to impede cell-cycle progression. All three inhibitors contain a ~60-amino-acid domain of similarity (39% to 47% sequence identity) in the N-terminal region and a nuclear localization signal near the carboxy terminus (Fig. 1). Outside these regions, however, these three proteins show no resemblance except for a short segment of similarity between p27 and p57 near their C termini. These differences suggested that the N-terminal domains of these inhibitors might contain a shared function, whereas the C-terminal regions might contain unique activities.

Previous work showed that a 52-amino-acid fragment of the p27 conserved N-terminal domain inhibits cyclin E-cdk2, albeit with only 5% of the potency of full-length p27 (ref. 5). To examine further the activity of this domain, we prepared a panel of recombinant p21, p27 and p57 derivatives (Fig. 1), all with a C-

terminal hexahistidine sequence for easy retrieval. These products were incubated with cyclin E-cdk2 complexes, the mixtures bound to nickel-Sepharose beads to retrieve inhibitor-bound components, and the bound material eluted, electroblotted and probed with cyclin E antibody. This assay should reflect binding to cyclin E-cdk2 complexes because the inhibitors bind to these complexes but not to the isolated components^{2,6}. When produced as truncated proteins or as part of chimaeric p21/p27 molecules, the N-terminal domains of human p21 (residues 1–75; N-p21 construct) and p27 (residues 1–86; N-p27 construct) had essentially full cyclin E-cdk2 binding activity as compared with the full-length proteins, whereas the C-terminal domains of p21 and p27 were inactive in this assay (Fig. 2a). The binding activity of p27 was decreased by deletion of 42 N-terminal amino acids and by lysine mutation of two residues, F33 and G34, that are conserved in p21 and p57 (Fig. 2a).

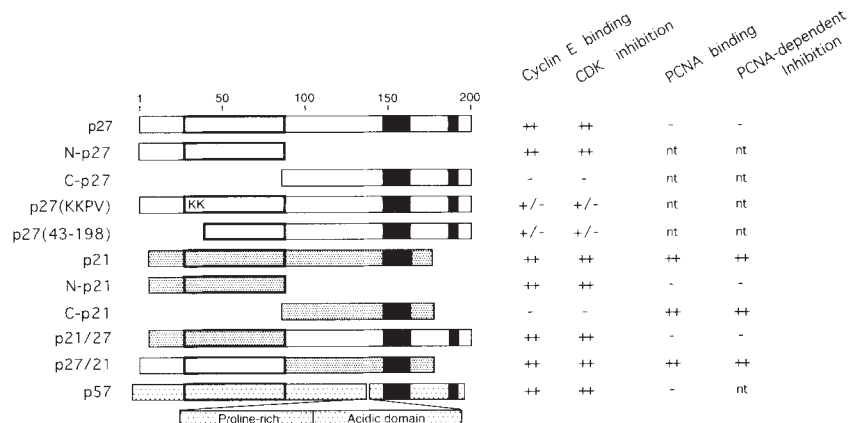
The CDK inhibitory activity of all these proteins paralleled their cyclin E-cdk2 binding activity. Using a glutathione *S*-transferase (GST)-Rb fusion protein as a kinase substrate, the isolated N-terminal domains of p21 and p27 inhibited cyclin E-cdk2 with a potency (ID₅₀ ~1 nM) similar to that of full-length p21 and p27 (Fig. 2b). The p27 mutants had only residual activity, whereas the isolated C-terminal domains of p21 and p27 were essentially inactive (Fig. 2b). Thus the N-terminal domains of p21 and p27 (and presumably p57) contain the CDK binding and inhibitory activities of these proteins.

To locate a PCNA interaction domain, the various histidine-tagged proteins were incubated with ³H-PCNA, retrieved with nickel-Sepharose and counted. PCNA binding activity was present in p21 but not in p27 or p57 (Fig. 2c). Furthermore, the isolated C-terminal domain of p21 bound PCNA, as did a chimaera containing this domain fused to the p27 N-terminal domain (Fig. 2c). The N-terminal domain of p21 alone or fused to the C-terminal domain of p27 did not bind PCNA. ³H-PCNA was coprecipitated with cyclin E-cdk2 complexes when the two were co-incubated in the presence of p21 but not in its absence (Fig. 2d), consistent with the ability of p21 simultaneously to bind PCNA and cyclin E-cdk2.

The PCNA-inhibitory activity of these proteins was determined by measuring their effect on PCNA-dependent DNA elongation by polymerase δ holoenzyme. A labelled primer on M13 DNA was elongated up to 7 kilobases (kb) by polymerase δ in a PCNA-dependent manner, and this reaction was inhibited by p21 (Fig. 2e), as previously reported^{9,10}. The C-terminal domain of p21 alone or as part of the p27/p21 chimaera inhibited this reaction whereas full-length p27, the N-terminal domain of p21 alone or the p21/p27 chimaera did not (Fig. 2d). Therefore the

FIG. 1 Schematic representation of p21^{Cip1/WAF1} (dark shading), p27^{Kip1} (open) and p57^{Kip2} (light shading) proteins, and summary of biochemical properties. The N-terminal region of similarity (thick box), nuclear localization sequence (closed box), C-terminal homology segment (small filled box), and the p57 proline-rich and acidic domains, are indicated. Scale indicates number of amino-acid residues. Biochemical activities summarize results shown in Fig. 2. The percentage activity of each protein relative to that of the corresponding full-length product is indicated by ++ (>50% activity), +/- (1–10%) or - (<1%); nt indicates not tested.

METHODS. Complementary DNAs encoding human p21 (refs 1, 3), human p27 (ref. 5), mouse p57 (ref. 7), the p21 N-terminal domain (codons 1–75) or C-terminal domain (codons 76–164), the p27 N-terminal domain (codons 1–86) or C-terminal domain (codons 87–198), a truncated p27 (codons 43–198), the F33K/G34K p27 double mutant, and the indicated chimaeras were generated by polymerase chain reaction (PCR) using appropriate templates and primers⁵. The



cDNAs were subcloned between *NheI* and *XhoI* sites in plasmid pET21a (Novagen), generating bacterial expression vectors encoding products with a C-terminal hexahistidine sequence. Expression and purification of recombinant proteins were done as described⁵.

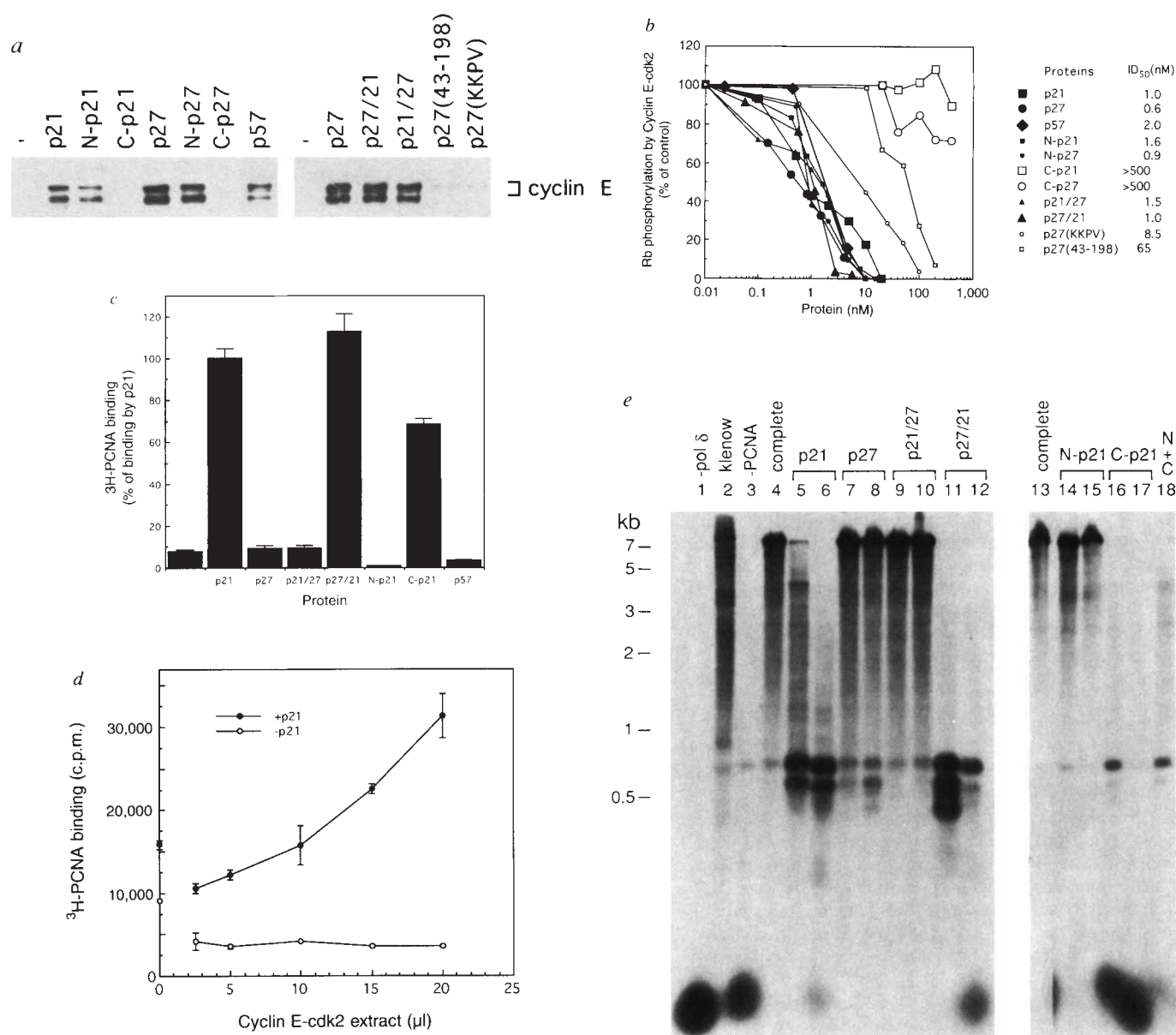


FIG. 2 *In vitro* activities of p21, p27 and p57 products. **a**, Cyclin E-cdk2 binding activity of the indicated products as determined by western immunoblot analysis of coprecipitating cyclin E. **b**, Inhibition of cyclin E-cdk2 Rb kinase activity by the indicated products. **c**, ³H-PCNA binding activity of various products relative to the binding activity of p21. **d**, p21-mediated association of PCNA with cyclin E-cdk2 as determined by coprecipitation of ³H-PCNA with cyclin E-cdk2 in the presence of p21. **e**, Effect of various p21 and p27 derivatives on the elongation of labelled singly primed DNA by polymerase δ holoenzyme. Reaction mixtures containing polymerase δ holoenzyme components with 0.17 μ M PCNA (as monomer), except in -pol δ and -PCNA lanes, and 5.4 fmol of labelled primed M13 DNA in the presence or absence of p21 and p27 products. **METHODS.** **a**, Cyclin E-cdk2 binding activity was determined by incubating insect cell lysates containing baculoviral-expressed cyclin E-cdk2 (ref. 5) with 8 nM of the indicated hexahistidine-tagged proteins in kinase buffer (20 mM Tris-HCl, pH 7.5, 7.5 mM MgCl₂, and 30 μ M ATP). Proteins were retrieved with Ni-NTA agarose (QIAGEN), resolved by SDS-PAGE (polyacrylamide gel electrophoresis), electroblotted and probed with human cyclin E antibody (Upstate Biotechnology Inc.) and ECL (Amersham). **b**, CDK inhibitory activity was determined by incubating cyclin E-cdk2 with the indicated proteins in kinase buffer containing 30 μ M [γ -³²P]-ATP and a GST-Rb fusion protein bound to glutathione-agarose beads (QIAGEN)¹³. Radioactivity incorporated into GST-Rb was quantified⁵ and plotted as percentage relative to controls without

inhibitors. **c**, PCNA binding activity was determined by incubating 72 ng of ³H-PCNA (6×10^2 c.p.m. ng⁻¹)¹⁴ with 750 nM of recombinant inhibitors in 50 μ l of binding buffer (20 mM Tris-HCl, pH 7.4, 100 mM NaCl, 1 mM EDTA, 0.1% NP40) at 37 °C for 5 min. Complexes were isolated with Ni-NTA agarose and the associated radioactivity counted and expressed as percentage relative to c.p.m. bound to full-length p21. **d**, p21-mediated association of PCNA with cyclin E-cdk2 was determined by incubating 180 ng of ³H-PCNA with increasing amounts of insect cell extract (16 mg per ml total protein) containing cyclin E-cdk2 (the amounts of cyclin E-cdk2 are limiting compared to that of PCNA or p21, even for the highest concentration) in 50 μ l of binding buffer at 37 °C for 15 min in the presence or absence of 40 nM recombinant p21. Cdk2-associated proteins were recovered by immunoprecipitation via a HA epitope tagged to the C terminus of cdk2 (ref. 15), and the associated radioactivity was counted. **e**, DNA replication assay was performed as previously described¹⁰. The following additions were made: lanes 5 and 6, 0.17 and 0.51 μ M p21; lanes 7 and 8, 0.26 and 0.78 μ M p27; lanes 9 and 10, 0.26 and 0.78 μ M p21/27; lanes 11 and 12, 0.32 and 0.96 μ M p27/21; lanes 14 and 15, 0.7 and 2.1 μ M N-p21; lanes 16 and 17, 0.48 and 1.44 μ M C-p21; lanes 18, 0.7 μ M of N-p21 plus 0.48 μ M C-p21. Lanes 4 and 13, containing the pol- δ holoenzyme and PCNA, were performed in different experiments, as were results presented in lanes 1-12 and 13-18. Lane 2, contained klenow (0.5 unit) in place of the pol- δ holoenzyme. Reaction products were resolved on alkaline agarose gels¹⁰ and autoradiographed.

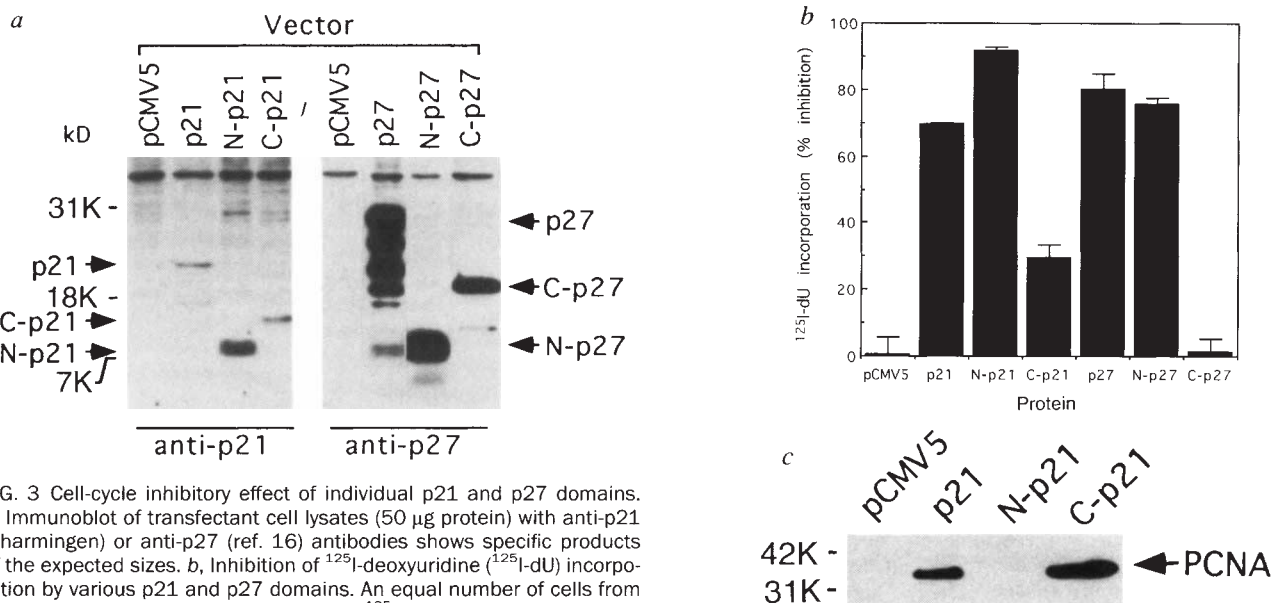


FIG. 3 Cell-cycle inhibitory effect of individual p21 and p27 domains. **a**, Immunoblot of transfectant cell lysates (50 μ g protein) with anti-p21 (Pharmingen) or anti-p27 (ref. 16) antibodies shows specific products of the expected sizes. **b**, Inhibition of [125 I]-deoxyuridine (125 I-dU) incorporation by various p21 and p27 domains. An equal number of cells from each transfection were incubated with [125 I]-dU for 2 hours, and the incorporated radioactivity counted¹⁷. Results are expressed as percentage inhibition of [125 I]-dU incorporation relative to cultures transfected with empty vector. Data are averages of triplicate determinations $\pm$ s.d. **c**, Anti-p21 immunoprecipitates from the indicated transfectants were resolved on SDS-PAGE, electroblotted and the blots probed with PCNA antibody (Santa Cruz Biotechnology). METHODS. cDNAs encoding the full-length p21, p27 or their C-terminal

PCNA binding and inhibitory activities are unique to p21, map to the C-terminal domain of this protein, and do not require the presence of a CDK inhibitory domain.

The presence of CDK and PCNA inhibitory activities in p21 has suggested^{9,10} that this protein may prevent execution of S phase in two ways: (1) by inhibiting G1 CDKs, thus impeding G1 to S transition, and (2) by inhibiting PCNA, which would directly interfere with DNA replication. The ability to separate physically these two activities allowed us to test this hypothesis by determining whether each active domain could interfere with cell-cycle progression when separately transfected into the mink lung epithelial cell line R-1B/L17 using transfection conditions that allow most of these cells to take up and transiently express transfected plasmids¹¹. To ensure nuclear localization of the isolated p21 and p27 N-terminal domains in transfected cells, these constructs were engineered with the simian virus 40 (SV40) T-antigen nuclear localization signal (PKKKRKV)¹² at their C termini. Immunoblotting of transfectant cell lysates with anti-p21 or anti-p27 antibodies demonstrated overexpression of all transfected constructs (Fig. 3a).

Transient transfection of p21, p27 or p57 into R-1B/L17 cells causes strong inhibition of DNA synthesis (as measured by [125 I]-deoxyuridine incorporation)⁵ (Fig. 3b) and accumulation of cells with a G1 DNA content⁵. Transfection of p21 or p27 N-terminal domains alone inhibited DNA synthesis as extensively as did transfection of the full-length proteins (Fig. 3b). Transfection of the p21 C-terminal domain alone was also inhibitory, the effect ranging between 25% and 40% inhibition ($n = 10$) (Fig. 3b). This effect was specific because the C-p27 protein did not have inhibitory activity (Fig. 3b), despite being expressed at a high level (Fig. 3a). Precipitation of transfected cell lysates with p21 antibodies followed by immunoblotting of the precipitates with anti-PCNA showed that the C-p21 protein binds PCNA *in vivo*, suggesting that the inhibitory effect of C-p21 resulted from its ability to bind PCNA (Fig. 3c). Quantitative western immunoblot assays indicated that the level of PCNA in R-1B/L17 cells was 0.23 pg per cell, of which only 25% to 30% was bound by transfected C-p21 that reached a level of 0.02 pg per cell. Thus

and N-terminal domains, all preceded by ~80 base pairs (bp) of the corresponding 5' UTR, were subcloned between the *Kpn*I and *Bam*HI sites in pCMV5 (ref. 11). The N-terminal domain constructs were engineered to encode a SV40 T-antigen nuclear localization signal PKKKRKV¹² at the C terminus. The resulting vectors were transiently transfected into R-1B/L17 cells^{5,11}, and analyses were done 48 hours post-transfection.

the incomplete inhibition of DNA synthesis by C-p21 in these transfectants may be due to incomplete titration of PCNA.

We have shown that members of the p21/p27 family contain an N-terminal CDK inhibitory domain that can inhibit cell-cycle progression independently of the C-terminal domain. By localizing the PCNA inhibitory function to the C-terminal domain of p21, and showing that this domain has antiproliferative activity as well, the evidence suggests that p21 can inhibit cell-cycle progression by two independent mechanisms. Both mechanisms are stoichiometric, their effectiveness in cell-cycle inhibition being determined by the relative abundance of their target CDKs and PCNA. The structural divergence of the p27 and p57 C-terminal domains suggests that they might also have unique regulatory functions.

Note added in proof: Warbrick *et al.*¹⁸ and Chen *et al.*¹⁹ have also identified the fits of CDK and PCNA binding in p21. $\square$

Received 23 January; accepted 16 March 1995.

1. El-Deiry, W. S. *et al.* *Cell* **75**, 817–825 (1993).
2. Gu, Y., Turck, C. W. & Morgan, D. O. *Nature* **366**, 707–710 (1993).
3. Harper, J. W., Adami, G. R., Wei, N., Keyomarsi, K. & Elledge, S. J. *Cell* **75**, 805–816 (1993).
4. Xiong, Y. *et al.* *Nature* **366**, 701–704 (1993).
5. Polyak, K. *et al.* *Cell* **78**, 59–66 (1994).
6. Toyoshima, H. & Hunter, T. *Cell* **78**, 67–74 (1994).
7. Lee, M.-H., Reynisdóttir, I. & Massagué, J. *Genes Dev.* **9**, 639–649 (1995).
8. Matsuo, S. *et al.* *Genes Dev.* **9**, 650–662 (1995).
9. Waga, S., Hannon, G. J., Beach, D. & Stillman, B. *Nature* **369**, 574–578 (1994).
10. Flores-Rozas, H. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **91**, 8655–8659 (1994).
11. Attisano, L. *et al.* *Cell* **75**, 671–680 (1993).
12. Kaideron, D., Roberts, B. L., Richardson, W. D. & Smith, A. E. *Cell* **39**, 499–509 (1984).
13. Matsushime, H. *et al.* *Cell* **71**, 323–334 (1992).
14. Kelman, Z., Naktinis, V. & O'Donnell, M. *Meth. Enzym.* (in the press).
15. Desai, D., Gu, Y. & Morgan, D. O. *Molec. Biol. Cell* **3**, 571–582 (1992).
16. Kato, J., Matsuo, K., Polyak, K., Massagué, J. & Sherr, C. J. *Cell* **79**, 487–496 (1994).
17. Laiho, M., DeCaprio, J. A., Ludlow, J. W., Livingston, D. M. & Massagué, J. *Cell* **62**, 175–185 (1990).
18. Warbrick, E., Lane, D. P., Glover, D. M. & Cox, L. S. *Curr. Biol.* **5**, 275–282 (1995).
19. Chen, J., Jackson, P. K., Kirschner, M. W. & Dutta, A. *Nature* **374**, 386–388 (1995).

ACKNOWLEDGEMENTS. We thank F. B. Dean, H. Flores-Rozas, Z. Kelman, M.-H. Lee, D. O. Morgan, M. O'Donnell, Z. Q. Pan and K. Polyak for reagents and advice, and B. Stillman for helpful discussions. This work was supported by a grant from the NIH to J.H. and by the Howard Hughes Medical Institute. J.H. is a professor of the American Cancer Society. Y.L. is a research associate and J.M. an investigator of the Howard Hughes Medical Institute.

Control of transcription by Krüppel through interactions with TFIIB and TFIIE β

Frank Sauer^{*†}, Joseph D. Fondell[‡],
Yoshiaki Ohkuma[‡], Robert G. Roeder[‡]
& Herbert Jäckle^{*}

^{*} Abteilung Molekulare Entwicklungsbiologie, Max-Planck-Institut für Biophysikalische Chemie, Am Fassberg, D-37077 Göttingen, Germany

[‡] Laboratory of Biochemistry and Molecular Biology,

The Rockefeller University, New York, New York 10021, USA

[†] Present address: Department of Molecular and Cell Biology, Howard Hughes Medical Institute, University of California, Berkeley, California 94720, USA

THE zinc-finger protein Krüppel (Kr)¹ is an integral part of the *Drosophila* segmentation gene cascade² and is essential in organogenesis during later embryonic development³. In tissue culture, Kr regulates transcription⁴⁻⁹. Monomeric Kr can act as a transcriptional activator, whereas Kr dimers formed at high concentrations cause repression⁶. Here we show that Kr-dependent control of transcription involves functional interactions with components of the basal RNA polymerase II transcription machinery, which includes the initiation factors TFIIA, B, E, F, H and I (refs 10, 11) as well as the TATA-binding protein (TBP) and TBP-associated factors (TAFs) contained in the multisubunit TFIID (ref. 12). Our results indicate that when acting from a site close to a basal promoter, monomeric Kr interacts with TFIIB to activate transcription, whereas an interaction of the Kr dimer with TFIIE β , a subunit of TFIIE, results in transcriptional repression.

To determine whether Kr can associate with the basal transcription machinery, we did coimmunoprecipitation experiments with Kr and recombinant initiation factors. The results shown in Fig. 1 indicate that Kr can associate with recombinant TFIIB¹³ and TFIIE β ¹⁴, but not with TFIIE α and TBP (Fig. 1a). The association with TFIIB and TFIIE β requires the presence of Kr-binding sites (Fig. 1a, b), suggesting that only DNA-bound Kr is capable of the observed interactions.

We analysed Kr-deletion mutants (Fig. 2a) for their ability to associate with TFIIB and TFIIE β . Mutant KRC-64 can bind to TFIIB (Fig. 2b), indicating that sequences containing the carboxy-terminal repressor and homodimerization domain of Kr (ref. 6) are dispensible for the association with TFIIB. Because

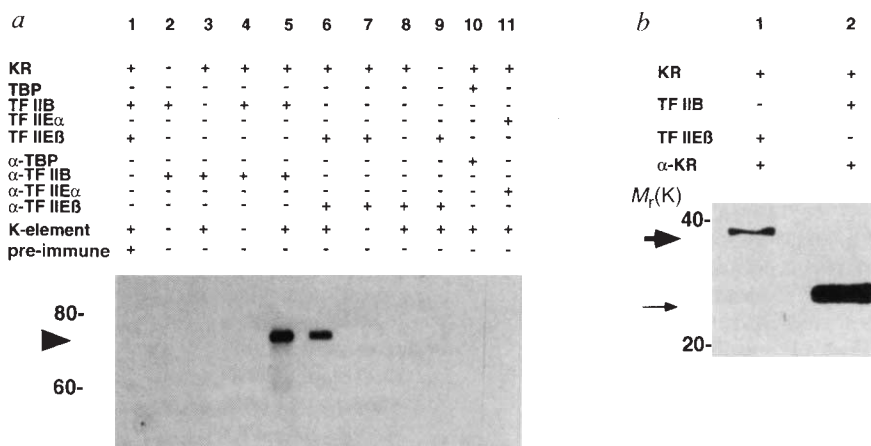
KRC-64 is incapable of homodimerization⁶, this result also suggests that the Kr monomer⁶ interacts with TFIIB to activate transcription. KRN-116 lacks the amino-terminal repressor domain⁹ and the N-terminal part of the activation domain which is located between amino acids 95 and 176 (ref. 15) (Fig. 2a). KRN-210 lacks the N-terminal repressor domain as well as the activation domain (Fig. 2a). Both KRN-116 and KRN-210 failed to associate with TFIIB (Fig. 2b, c). However, TFIIB could associate with KR95-176 (Fig. 2c; lane 3), indicating that the region of Kr required for activation is sufficient for the interaction with TFIIB. Because KR95-176 lacks the DNA-binding domain (Fig. 2a) and Kr-binding sites are not required for the interaction observed between KRC-64 and TFIIB (Fig. 2c), full-size Kr is likely to undergo a conformational change upon DNA binding that makes the activation domain accessible for the interaction with TFIIB.

TFIIE β could not associate with KRC-64 or KRN-116 (Fig. 2b; lanes 2 and 4). Thus, sequences contained within both the N- and C-terminal regions of Kr are required for this interaction. Because KRC-64 lacks the homodimerization domain⁶, we wondered whether TFIIE β did not interact with KRC-64 because it fails to form dimers or because TFIIE β normally binds to the KR C-terminal sequences not present in KRC-64. To distinguish between these possibilities, we examined the binding of TFIIB and TFIIE β to dimeric Kr. Interaction analyses with crosslinked proteins (Fig. 3a, b) revealed that the Kr dimer interacts with TFIIE β but not with TFIIB (Fig. 3c). Furthermore, TFIIE β can associate with the crosslinked Kr C-terminal portion (Fig. 3d) represented by C64KR (Fig. 2a). Because KRN-116 was unable to coimmunoprecipitate with TFIIE β (Fig. 2b, lane 2) and because TFIIE β could not associate with a peptide encompassing the N-terminal 116 amino acids (data not shown), we suspect that crosslinking of C64KR provides the stabilization of the dimer normally achieved through N-terminal Kr sequences.

To see whether interactions of Kr with TFIIB and TFIIE β have functional consequences, we assayed transcriptional regulation by Kr *in vitro*¹⁶. Kr can activate or repress transcription in a concentration-dependent manner (Fig. 4a). The monomeric KRC-64 activated transcription (Fig. 4b, d, g), whereas the crosslinked Kr dimer repressed transcription (Fig. 4c, e). Activation by KRC-64 was inhibited in the presence of the anti-TFIIB antibodies (Fig. 4d, g). Conversely, repression by the crosslinked Kr dimer was inhibited by anti-TFIIE β antibodies (Fig. 4e). Although rendering the system non-responsive to Kr (Fig.

FIG. 1 Kr associates with TFIIB and TFIIE β *in vitro*. a, Coimmunoprecipitation followed by SDS-PAGE using *in vitro* translated ³⁵S-labelled Kr (arrowhead), unlabelled recombinant TBP, TFIIB, TFIIE α or TFIIE β and the corresponding antibodies (α -TBP, α -TFIIB, α -TFIIE α and α -TFIIE β). Note that *in vitro* association of Kr with TFIIB or TFIIE β requires the presence of oligonucleotides containing a Kr-binding site of the sequence -AAAAGGGTTAA- (K-element; compare lanes 4, 5 and 6, 7). b, Coimmunoprecipitation followed by SDS-PAGE using ³⁵S-labelled TFIIB (small arrow) or TFIIE β (large arrow), unlabelled Kr and antibodies directed against Kr (α -Kr). The K-element was included in the experiments. Note that neither TFIIB nor TFIIE β can bind to the K-element as revealed by gel mobility shift assays and immunoprecipitation experiments involving the K-element (data not shown; see also Figs 2b, c and 3c, d). *M_r* markers are indicated at the left; 'pre-immune' refers to the pre-immune sera.

METHODS. Kr was *in vitro* transcribed and translated as described⁶. Unlabelled and ³⁵S-methionine-labelled recombinant TFIIB and TFIIE β were produced²⁶ and purified under native conditions²⁷. Coimmunoprecipitation assays were assayed as described⁶. They contained 10 μ l *in*



in vitro translated Kr, 100 ng affinity-purified TFIIB or TFIIE β in 30 μ l 1 $\times$ binding buffer (BB; 25 mM HEPES pH 7.9, 100 mM KCl, 20% glycerol, 1 mg ml⁻¹ BSA, 1 mM ZnSO₄, 1 mM spermine) and optionally 10 ng of gel-purified K-element oligonucleotide containing a single Kr target sequence⁵.

4d, e, g), neither of the antibodies interfered with basal transcription (Fig. 4f; see also Fig. 4d, e, g), presumably because the antibodies recognize only those domains of TFIIB and TFIIE β required for functions which are not apparent from the *in vitro* assay. Addition of recombinant TFIIE β partially restored Kr-dependent repression (Fig. 4e, lanes 6–8). Conversely, the addition of recombinant TFIIB partially restored activation (Fig. 4g; lanes 6–8). Thus, promoter complexes containing TFIIB or TFIIE β associated with Kr may be more stable than promoter complexes with antibody-bound TFIIB or TFIIE β .

Our results suggest that the association of Kr and TFIIE β causes transcriptional repression, whereas the association of Kr

with TFIIB causes activation. Interactions between activators and TFIIB help the recruitment of other basal initiation factors^{16–21}, but an interaction between a repressor and a TFIIE subunit has not yet been described. Because the role of TFIIE in initiation complex formation and function involves the modulation of ATPase, C-terminal domain kinase and helicase activities of TFIIE^{22–24}, the repression-related interaction of Kr with TFIIE β may reflect the inhibition of one of these

FIG. 2 TFIIB and TFIIE β interact with distinct regions of Kr. **a**, Schematic representation of Kr and various Kr deletion mutants. Wild-type Kr is shown on top (KR). It contains two repressor domains^{5,6,9} (rd1 and rd2), an N-terminal activator domain^{5,15} (black bar), the C-terminal dimerization domain⁵ (dd), the DNA-binding zinc-finger region¹ (open bar; Cys2/His2) and a nuclear translocation signal (NL). Deletion mutants KRN-116, KRN-210, KR95–176, KRC-64 and C64KR are shown below; Met refers to the experimentally induced translation start sites. **b**, Coimmunoprecipitation followed by SDS-PAGE separation using ³⁵S-labelled truncated versions of Kr. TFIIB associates with KRC-64 (lane 3) but not with KRN-116 (lane 1) or KRN-210 (c). Thus, N-terminal sequences are necessary for the Kr-TFIIB interaction. TFIIE β does not interact with KRN-116 (lane 2), KRC-64 (lane 4) and KRN-210 (data not shown). Thus, both C-terminal and N-terminal sequences of Kr are necessary for the TFIIE β interaction. *In vitro* translated KRN-116 (arrowhead) and KRC-64 (arrow) are shown in lanes 5 and 6, respectively. Note that the K-element was present in all experiments shown (see also Fig. 1). **c**, The Kr sequence interval 95–176 is sufficient for the interaction with TFIIB. Coimmunoprecipitation experiments (lanes 2–5) followed by SDS-PAGE⁶ were done using the *in vitro* translated unlabelled KRC-64, KR95–176, KRN-116 or KRN-210 together with ³⁵S-labelled recombinant TFIIB (shown in lane 1; control). All reactions lack the K-element. Thus, KRC-64 (lane 2) and KR95–176 (lane 3), which contains the N-terminal activator domain of Kr (ref. 15) do not require the K-element for an association with TFIIB, whereas the presence of Kr-binding sites is absolutely essential for the association of full-size Kr and TFIIB (Fig. 1). Abbreviations as for Fig. 1. **METHODS**. The proteins were produced and purified as described in Fig. 1. The reactions contained 10 μ l reticulocyte lysate programmed with Kr mRNA, 50 ng of ³⁵S-labelled recombinant TFIIB and 10 ng of K-element. Coimmunoprecipitations using anti-Kr antibodies were as described in Fig. 1 using KRN-116, KRC-64 or KR95–176 instead of Kr. The *in vitro* transcription and translation of KRN-116 and KRC-64 has previously been described⁶. To generate KR95–176, the 240-bp *Ball-Nhl* Kr fragment¹ was fused to linker sequences that create a start codon at position 95 before insertion into pBluescript (Stratagene).

FIG. 3 TFIIE β interacts with the Kr dimer. Purification and SDS-PAGE analysis of glutaraldehyde-crosslinked Kr dimers (a) and C64KR dimers containing the Kr C-terminus (b). ³⁵S-methionine-labelled recombinant Kr (a) or C64KR (b) were removed from crude bacterial lysates (a, b, lanes 1) by Ni²⁺-affinity chromatography. Purified Kr (a, arrow in lane 2) was crosslinked by glutaraldehyde and dimers (a, c, arrowhead) were isolated (a, lane 3). Correspondingly, purified C64KR was crosslinked by glutaraldehyde (lane 2 in b; note unreacted C64KR marked by arrow) and dimers (b, c, arrowhead) were isolated (b, lane 3). Crosslinked Kr or C64KR dimers were used for coimmunoprecipitation experiments (c, d) involving recombinant TFIIB or TFIIE β and the corresponding antibodies (α -TFIIB, α -TFIIE β). The reactions in c included 10 ng of K-element which was absent from reactions in d. Note that both dimers associate with TFIIE β (c, d, lanes 1) but not with TFIIB (c, d, lanes 2). Thus, TFIIE β interacts with the Kr C-terminal region when dimerized. This suggests that the C64KR domain is directly required for the interaction of Kr with TFIIE β . Note that the Kr dimers, but not the C64KR dimers, require the presence of the K-element for their association with TFIIE β . *M_r* markers are indicated on the left in a–d; abbreviations as for Fig. 1.

METHODS. For expression of histidine-tagged Kr derivatives in *Escherichia coli*, the 1.8 kb *NdeI*–*EcoRI* fragment or the 0.8-kb *Sall* fragment of Kr cDNA that derived from pBlueKr⁶ was fused in-frame to the histidine-tag coding sequence of the plasmid pRSet (Invitrogen). Fusion genes were expressed in the presence of ³⁵S-labelled methionine as described in Fig. 1. The purified proteins were crosslinked as previously described⁶. The monomeric form of C64KR was excluded from the dimeric form by cut-off centrifugation using a Centricon filter spin column (Amicon company). Coimmunoprecipitations were as described in Fig. 1 using 200 ng of the crosslinked Kr or C64KR dimers and 100 ng TFIIE β . SDS-PAGE was done as described in Fig. 1.

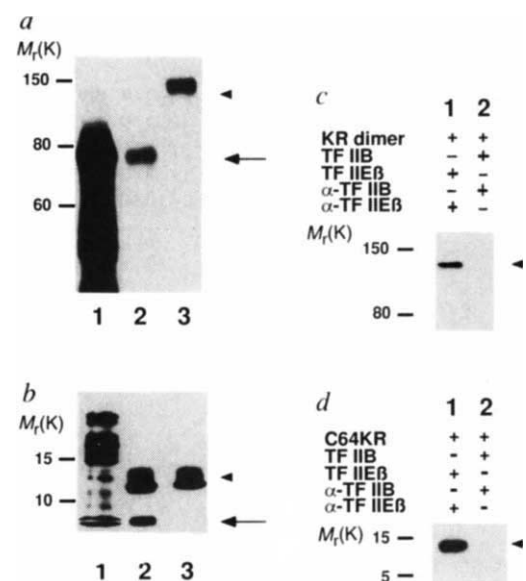
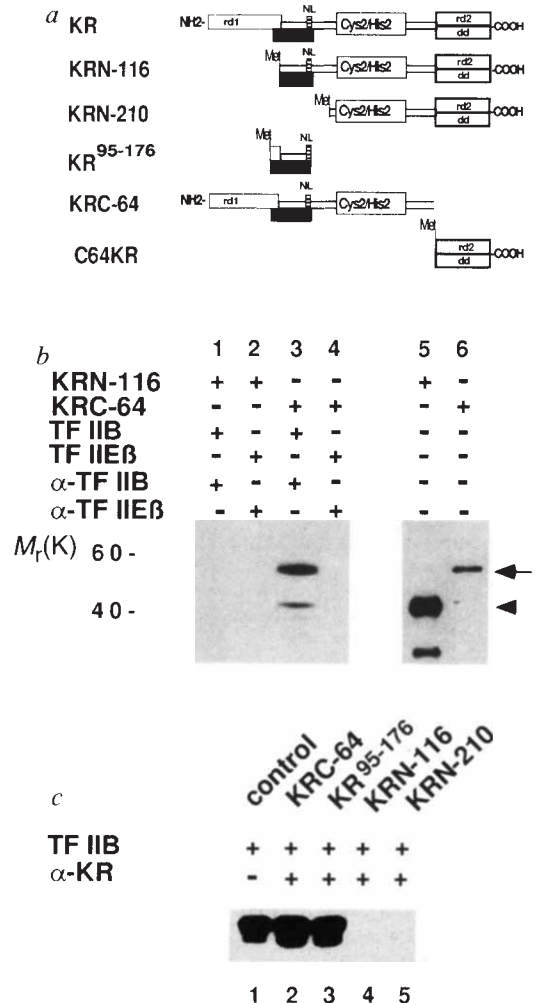
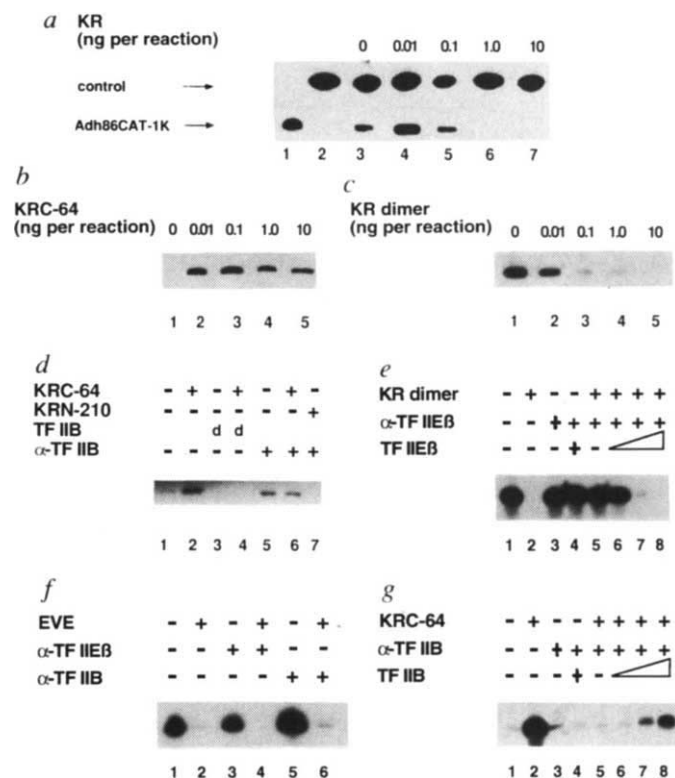


FIG. 4 Interaction with TFIIB or TFIIE β is required for Kr-dependent activation or repression of *in vitro* transcription. **a**, Concentration-dependent activation and repression of transcription by Kr. Kr activates (lane 4) or represses (lanes 5–7) Adh86CAT-1K^{5,6} transcription (lower lane). Note that Adh86CAT-1K contains a single Kr-binding site, the *cis*-acting K-element and that no Kr regulatory response was observed in the absence of the *cis*-acting K-element (upper lane; 'control', see Methods). **b**, Adh86CAT-1K repression by the crosslinked Kr dimer; the different basal level of transcription is due to the different amount of reporter gene DNA (200 ng per reaction). **c**, α -TFIIB antibodies inhibit KRC-64-dependent activation. Transcription of Adh86CAT-1K (50 ng per reaction) in the absence (lanes 1, 3, 5, 7) or presence of 0.01 ng KRC-64 (lanes 2, 4, 6). KRC-64-dependent activation (lane 2) was inhibited when TFIIB was removed from the extracts (lane 4; 'd' refers to depletion) or anti-TFIIB antibodies (α -TFIIB) were added (lane 6). The latter affected neither KRN-210-dependent repression (compare lanes 5 and 7) nor the basal level of Adh86CAT-1K transcription (lanes 5, 6; see also e–g). **d**, Anti-TFIIE β antibodies (α -TFIIE β) prevent Kr dimer-dependent repression in a reversible manner. Adh86CAT-1K (200 ng per reaction) transcription in the absence (lanes 1, 3, 4) or presence (lanes 2, 5–8) of crosslinked Kr dimer (1.0 ng per reaction). Kr dimer-dependent repression (lane 2) was inhibited by α -TFIIE β (lane 5) which did not affect basal transcription (compare lanes 1, 3 and 4; see also f). Addition of recombinant TFIIE β (10 ng in lane 4; 1, 5 and 10 ng in lanes 6–8, respectively) enables the crosslinked Kr dimer to repress transcription in the presence of α -TFIIE β (lanes 6–8). **e**, In the presence of either α -TFIIE β (lane 4) or α -TFIIB (lane 6), repression of the reporter gene *Even-skipped* (Eve) can be achieved by 50 ng of the transcriptional repressor encoded by the pair-rule gene *even-skipped* (Eve). **f**, α -TFIIB prevents KRC-64-dependent activation in a reversible manner. Adh86CAT-1K (50 ng per reaction) transcription in the absence (lanes 1, 3, 4) or the presence of KRC-64 (0.1 ng per reaction; lanes 2, 6–8). KRC-64-dependent activation (lane 2) was inhibited by α -TFIIB (lane 5) which did not affect basal transcription (lane 3; also see f). Addition of recombinant TFIIB (7 ng in lane 4; 1, 4 and 7 ng in lanes 6–8, respectively) enables KRC-64 to activate transcription in the presence of α -TFIIB (lanes 6–8). Note different basal transcription levels in e and f because of different amounts of reporter gene DNA in the reactions. Abbreviations are as described in the legend of Fig. 1.

METHODS. Kr and KRC-64 (see Figs 1 and 2) were dialysed against BC50²³. Crosslinked Kr dimer was generated by glutaraldehyde⁶, separated by SDS-PAGE, recovered by electroelution and renatured by extensive dialysis against BC50 containing decreasing concentrations of SDS. *In vitro* transcription²⁸ was done with 50 μ g HeLa nuclear extract²⁹, 300 ng pKRCAT⁵ and the indicated amounts of pAdh86CAT-1K⁵. Transcripts were assayed by primer extension with a 5' ³²P-labelled oligonucleotide 5'-CAACGGTGTATATCCAGTG-3' complementary to the



coding region of the CAT gene (giving rise to a 126 nt reverse transcript); a 456-nt reverse transcript that derived from the K-element-less reporter plasmid pKRCAT⁵ was coassayed as an internal control (shown in a). Eve was expressed from plasmid pAReve and purified as described³⁰. For antibody treatment, nuclear extracts (1 mg) were preincubated with α -TFIIB (1.2 μ g of total protein) or α -TFIIE β (0.5 μ g of total protein) antisera for 4 h at 4 °C. For competition experiments, nuclear extracts were preincubated with antisera (1 h at 4 °C) before recombinant TFIIB or TFIIE β were added. Antibody concentrations that selectively inhibited Kr-dependent activation or repression were determined empirically (50 ng α -TFIIB; 20 ng α -TFIIE β in d–g) so that neither basal transcription nor control gene expression were significantly altered (see d–g). TFIIB-depleted nuclear extracts were obtained as described¹⁸.

functions or a destabilization of the initiation complex. Within this context it is interesting to note the control experiments (Fig. 4f) in which the repressor *even-skipped* (Eve) exerts repression in the presence of anti-TFIIE β -antibodies that otherwise prevent Kr-dependent transcriptional repression (Fig. 4e). This suggests that the mechanisms of repression by Kr and Eve are different.

The different interactions of Kr shown here may represent the molecular basis of Kr-dependent transcriptional regulation in *Drosophila* Schneider cells^{5,6}, which is mediated by a single Kr-binding site next to the basal promoter. When Kr acts from a different binding site in mammalian tissue culture cells, it

represses transcription by quenching through the N-terminal repressor domain⁹. Furthermore, Kr provides mainly, if not exclusively, repression through distant enhancers when acting within the segmentation gene cascade, possibly involving a quenching mechanism as well^{7,8,25}. Our findings may therefore apply to specific aspects of Kr-dependent transcriptional control at later stages of embryogenesis and/or might be restricted to specific Kr-binding sites. Given our limited knowledge concerning molecular mechanisms that govern developmental decisions and the various and mostly unknown Kr target genes, our findings provide a framework from which these questions can be addressed systematically. □

Received 22 September 1994; accepted 13 March 1995.

- Rosenberg, U. et al. *Nature* **319**, 336–339 (1986).
- Ingham, P. W. *Nature* **335**, 25–34 (1988).
- Schmucker, D., Taubert, H. & Jäckle, H. *Neuron* **5**, 1025–1039 (1992).
- Licht, J. D., Grossel, M. J., Figge, J. & Hansen, U. *Nature* **346**, 76–79 (1990).
- Sauer, F. & Jäckle, H. *Nature* **353**, 563–565 (1991).
- Sauer, F. & Jäckle, H. *Nature* **364**, 454–457 (1993).
- Zhuo, P. et al. *Genes Dev.* **5**, 254–264 (1991).
- Small, S., Kraut, R., Hoey, T., Warrior, R. & Levine, M. *Genes Dev.* **5**, 827–839 (1991).
- Licht, J. D., Ro, M., English, M. A., Grossel, M. J. & Hansen, U. *Proc. natn. Acad. Sci. U.S.A.* **90**, 11361–11365 (1993).
- Roeder, R. G. *Trends biochem. Sci.* **16**, 402–407 (1991).
- Zawel, L. & Reinberg, D. *Curr. Opin. Cell Biol.* **4**, 488–495 (1992).
- Hernandez, N. *Genes Dev.* **7**, 1291–1308 (1993).
- Malik, S. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 9553–9557 (1991).
- Sumimoto, H. et al. *Nature* **354**, 401–404 (1991).
- Sauer, F. thesis, Univ. Osnabrück, Germany (1993).
- Lin, Y. S. & Green, M. R. *Cell* **64**, 971–981 (1991).

- Choy, B. & Green, M. R. *Nature* **366**, 531–536 (1993).
- Kim, T. K. & Roeder, R. G. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Horikoshi, M. et al. *Cell* **54**, 665–669 (1988).
- Stringer, K. F., Ingles, C. J. & Greenblatt, J. *Nature* **345**, 783–786 (1990).
- Goodrich, J. A., Hoey, T., Thut, C. J., Amon, A. & Tjian, R. *Cell* **75**, 519–530 (1993).
- Lu, H., Zawel, L., Fischer, L., Egly, J.-M. & Reinberg, D. *Nature* **358**, 841–845 (1992).
- Ohkuma, Y. & Roeder, R. G. *Nature* **368**, 160–163 (1994).
- Goodrich, J. A. & Tjian, R. *Cell* **77**, 145–156 (1994).
- Stanojevic, D., Small, S. & Levine, M. *Science* **254**, 1385–1387 (1991).
- Studier, F. W. & Moffat, B. A. *J. molec. Biol.* **189**, 113–130 (1986).
- Hoffmann, A. & Roeder, R. G. *Nucleic Acids Res.* **19**, 6337–6338 (1991).
- Lillie, J. W. & Green, M. R. *Nature* **338**, 39–44 (1991).
- Dignam, J. D., Lebovitz, R. M. & Roeder, R. G. *Nucleic Acids Res.* **11**, 1475–1489 (1983).
- Han, K. & Manley, J. L. *Genes Dev.* **7**, 491–563 (1993).

ACKNOWLEDGEMENTS. We thank our colleagues in both laboratories for their various contributions. This work was supported in part by grants from the NIH and by general support from the Pew Trusts to the Rockefeller University (R.G.R.), and by the Max-Planck Society and the Fonds der Chemischen Industrie (H.J.). F.S. was supported by a Max-Planck fellowship; J.D.F. was supported by a postdoctoral fellowship from the NIH.

Microbiology and manufacturing

Adding culture to microbiology, product review presents a portable case for microbial decontamination, a HEPA-filtered infrared CO₂ incubator, a modular fermentor, and a family of X-ray detectors for electron microscopes.

THE Free Zone Plus six-litre cascade freeze-dry system offers lower condenser temperatures with environmentally safe refrigerants (*Reader Service No. 100*). It features two refrigeration systems, used in series, to lower condenser temperature to -84°C , allowing lyophilization of hard-to-freeze materials with lower eutectic points. It has automatic start-up to initiate the condenser cool-down and vacuum pull-down sequence. The LED allows the user to see both the vacuum reading and the condenser temperature at a glance. A power failure indicator illuminates to signal that power was interrupted.

The Decontakit is a portable case for microbial decontamination of laminar flow cabinets, biohazard safety cabinets, CO₂ incubators, animal storage cabinets and the like (*Reader Service No. 101*). This kit is manufactured by ESI Flufrance and is a complete formaldehyde-based decontamination system, which decontaminates a device, on site, without perturbation in the laboratory. It includes one fully automatic decontamination and neutralization system, one separable part with generators to be set inside the contaminated area, accessories and consumables for two cycles. The Decontakit operates automatically overnight — install before exiting the laboratory, unplug it upon returning and resume working. It is designed to remove any trace of formaldehyde, even odours.

■ Incubation

Don Whitley Scientific has developed its new anaerobic cabinet to the changing requirements of the microbiologist. (*Reader Service No. 102*). The redesigned size, shape, internal dimensions, angle of viewing panel, control layout and porthole detail are the result of careful ergonomic consideration. The current capacity of the

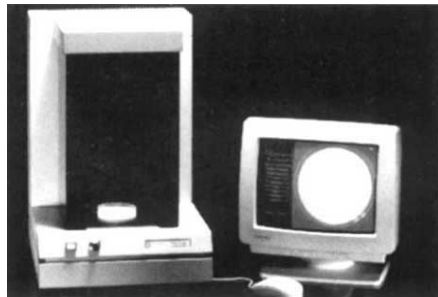
company's range of incubators is between 200 and 1,000 plates. This latest addition will accommodate up to 540 9-cm Petri dishes. All systems include 'three gas mixing' that allows lower running costs and 'anatox atmospheric scrubbing' to remove volatile fatty acids and prolong the life of the catalyst. The elongated shape of the porthole entry is designed to allow more comfort in use and greater freedom of movement inside the chamber. Each porthole also provides a means of transferring up to ten 9-cm dishes.

Forma Scientific has introduced the new Steri-cult 200 HEPA-filtered infrared CO₂ incubator (*Reader Service No. 103*). It has been designed to overcome the problems of contamination and desiccation encountered during incubation of microbiological cultures. One hundred per cent of recirculated air passes through the HEPA filter, ensuring constant decontamination, minimizing contamination and eliminating the need for other methods of heat disinfection. The company says that the temperature, humidity and CO₂ level can be accurately and independently controlled to suit user needs (temperature, up to 14°C above ambient to 45°C with an overall uniformity of ± 0.3 ; CO₂ ranges from 0 to 20 per cent; and humidity, ± 2.0 per cent with humidity reservoir of 3.75 litre). Two sizes are available.

■ Colony counting

A new high-throughput colony count-ing system that utilizes video and menu-driven computer technology to provide accurate counts of bacteria, cells and plaques on a wide variety of media has been introduced by Optomax (*Reader Service No. 104*). The Cardinal automatic colony counting system features a compact plate viewer for colonies on transparent and opaque nutrient media with software selectable bright- and dark-field illumination. The system provides a Windows environment for point-and-click operation, while creating a complete audit trail. It is capable of up to 600 plates per hour and typically performs measurements in 0.5 s with 0.15-mm resolution at full plate magnification. An option allows direct connection to a light microscope for counting silver grains, nuclei or fluorescing bacteria.

Protos Plus is a colony-counting system from Synoptic designed for rapid counting of colonies even on opaque media (*Reader Service No. 105*). This makes it suited to applications in the water and pharmaceutical industries. Protos Plus features a new processor that is stated to



Protos Plus counting system.

make all processing operations up to four times faster. The manufacturer states that features such as the automatic separator facility, which uses mathematical morphology to separate touching colonies for counting, operate seamlessly with other operations.

Celsis now offers the Scan 500 colony counter for enumeration of micro-organisms grown on agar plates (*Reader Service No. 106*). This counter utilizes a design to reduce operator fatigue, eliminate inaccuracies and improve the slow speed associated with manual counting of poured and spiral plates. In addition, the counter uses video camera technology that provides a consistent method of counting microbes, which helps standardize results while reducing handling time. The instrument is comprised of a charge-coupled device camera and pattern recognition software for accurate counting of difficult colonies; such as, spreaders, pinpoints and continuous colonies. The system is menu-driven and includes on-screen prompts to simplify the testing procedure for laboratory personnel.

■ Culturing

The new BioFlo 1000 from New Brunswick Scientific is designed to be an economical modular fermentor system that can be built up to suit most applications (*Reader Service No. 107*). This system can be used for experimental modelling set-ups such as environmental and bacterial growth monitoring. The researchers can identify relevant features and tailor the fermentor system to their budget without limiting the ability to expand. The modular design is also advan-



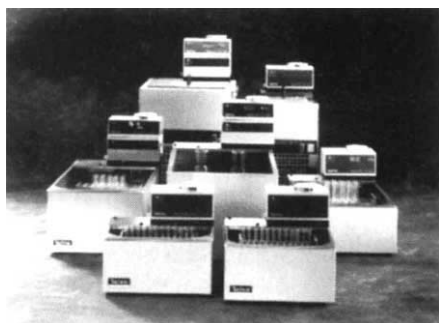
Don Whitley Scientific anaerobic cabinet.

PRODUCT REVIEW

tageous for multiple user situations — features can be adapted for particular purposes, for example, growth of bacterial, yeast and fungal cultures, with the facility to modify for shear-sensitive organisms.

Biomed Diagnostics has introduced the **InPouch TV subculture kit** designed to keep *Trichomonas vaginalis* stock cultures alive for more than one week (*Reader Service No. 108*). The kit uses a highly selective growth medium that is designed to support the slow growth of *T. vaginalis* over longer time periods. It also maintains optimum culture integrity by eliminating fungal and bacterial growth. This subculture kit is a companion to the **InPouch TV *T. vaginalis* diagnosis kit**, and is stated to have a sensitivity of 4 CFU ml⁻¹. The kit uses a two-chamber, high-oxygen barrier plastic pouch as an incubator and transport device. Viewing is made directly through the clear pouch, eliminating sample transfer and handler exposure as well as reducing hands-on time.

Techné's constant temperature water/oil baths offer the choice of 25 different combinations of bath and controller (*Reader Service No. 109*). There are models available with temperature ranges from



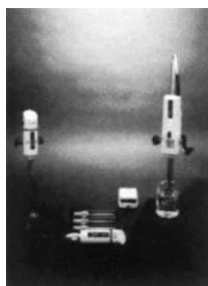
Techné's oil and water baths.

–35 °C to +200 °C and sizes from 8–48 litres. There are digital or analog displays available. Temperature control stability is good to ±0.005 °C, says Techné.

■ Miscellaneous

Unipath has launched Quanti-cult plus as part of the new Oxoid range of specialty products for quality control purposes (*Reader Service No. 110*). Each vial contains stabilized microorganisms and will deliver a specific range of colony-forming units (CFUs) following rapid and easy rehydration. This system enhances aseptic technique as ten 0.1-ml source inoculations of 10–100 CFUs are dispensed from a single 1.2-ml source without compromising the purity of the remaining units. It is designed to be cost-effective, to save time and to eliminate the need for the maintenance of stock cultures, lengthy growth periods and trial-and-error dilution of suspensions to achieve the desired count. The procedure requires no 'sharps', pipettes or forceps.

For handheld or post-mounted **homogenizers**, Omni presents the H series, tailored for specific laboratory applications (*Reader Service No. 111*). These feature powerful 30,000 r.p.m. motors to process samples ranging from 0.25- to 30-litre volumes. The µH model is a rechargeable, battery-operated unit weighing only 311 g. It is suited for repetitive handheld operation and is portable. The tissue homogenizer, TH, processes even fibrous samples, in times stated of less than 30 s. The general laboratory homogenizer is a versatile instrument offering an optional external speed control for remote or fume-hood processing, as well as a choice of 15 rotor stator accessories.



Omni homogenizer.

Princeton Gamma-Tech has recently announced the addition of JT Cool, with the added option for no-liquid nitrogen, to the Prism family of **X-ray detectors for electron microscopes** (*Reader Service No. 112*). This system provides stable, cryogenic cooling in situations where liquid nitrogen cannot be used. It provides resolution and light element sensitivity, including Be detection.

Precision Scientific offers a line of **benchtop autoclaves** designed for the sterilization of glassware and instruments as well as solid and liquid media (*Reader Service No. 113*). Intended for use in a wide range of disciplines, Napco autoclaves deliver sterilization temperatures as high as 132 °C at pressures up to 186 kPa. In addition, these instruments incorporate a continuous pressure purge system that automatically eliminates air pockets and provides uniform penetration of superheated steam to all areas of the sterilization/test chamber. Three models are available — all feature a drying element to remove residual moisture at the end of a cycle, door-mounted pressure and temperature gauges, an over-temperature safety thermostat and an over-pressure relief system. Each has a stainless steel sterilization/test chamber measuring 23 cm in diameter by 46 cm long. The model 9000D is a fast cool-down autoclave used for sterilizing instruments and solids. The 8000-DSE has a slow cool-down capability to prevent liquid boilover and can be used with both liquid and solid materials. The 8100-TD is modified with an environmental test chamber for use with electronics.

Manostat has announced the addition of a green pen to the black and red colony

marker pens it already offers (*Reader Service No. 114*). The Colony Counter pens have a replaceable long-life felt tip and may be used on glass or plastic Petri dishes. The use of three colours is designed to allow changes to be tracked more easily.

Over 200 items are featured in Sterilin's full-colour **catalogue of plasticware products** (*Reader Service No. 115*). Sterilin offers a wide choice of disposable plasticware, from Petri dishes, pipettes and specimen containers to special purpose items. □

These notes are compiled by Brendan Horton from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.

ADVERTISEMENTS

OLIGONUCLEOTIDES

Synthesised, purified and shipped in 3 days

£2.50 per base

(All inclusive)

0.2 µM scale (approx. 500 µg)

- Antisense (completely or partially sulphurised)
- Extensive range of labelling and modifications
- Full design and analysis facilities
- Gene cassettes, bulk order and contract discounts

Molecular Medicine Unit

King's College School of Medicine & Dentistry
Tel: 071 346 3126 Fax: 071-733 3877

Reader Service No. 10

DNA, PEPTIDE SYNTHESIS

PEPTIDE/PROTEIN SEQUENCING

Personalized Service
Guaranteed Quality

Peptides: \$26.99/ residue*

Oligos: \$3.00 / base

*minimum 100

DNA - PEPTIDES SHIPPED
WORLD WIDE DAILY
PROVEN PERFORMER
VETROGEN CORP
A Canadian Company

TEL: 1-519-652-3758

FAX: 1-519-652-3818

Reader Service No. 12

XX IDT
INTEGRATED DNA TECHNOLOGIES, INC.

Custom DNA Synthesis

\$1.95 / Base

200 nmole Scale

No Set-up Fee

Column Purified

24 Hour Shipping

Q.C. by PAGE on Every Oligo

E-mail: orders@bbs.idtdna.com

Fax: 1-319-645-2921

Phone: 1-800-328-2661

Major Credit Cards Accepted

Reader Service No. 3